



The fixed point property and a technique to harness double fixed point combinators

Giulio Manzonetto, Andrew Polonsky, Alexis Saurin, Jakob Grue Simonsen

► To cite this version:

Giulio Manzonetto, Andrew Polonsky, Alexis Saurin, Jakob Grue Simonsen. The fixed point property and a technique to harness double fixed point combinators. *Journal of Logic and Computation*, 2019, 29 (5), pp.831-880. 10.1093/logcom/exz013 . hal-02408192

HAL Id: hal-02408192

<https://hal.science/hal-02408192>

Submitted on 7 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The fixed point property and a technique to harness double fixed point combinators

GIULIO MANZONETTO, *Université Paris 13, Laboratoire d'Informatique de Paris-Nord, CNRS UMR 7030, France.*

ANDREW POLONSKY, *Department of Computer Science, Appalachian State University Boone, NC 28608, USA.*

ALEXIS SAURIN, *Université de Paris, IRIF, CNRS, F-75013 Paris, France.*

JAKOB GRUE SIMONSEN, *Department of Computer Science, University of Copenhagen (DIKU), Universitetsparken 1, 2100 Copenhagen Ø, Denmark.*

Abstract

The λ -calculus enjoys the property that each λ -term has at least one fixed point, which is due to the existence of a fixed point combinator. It is unknown whether it enjoys the ‘fixed point property’ stating that each λ -term has either one or infinitely many pairwise distinct fixed points. We show that the fixed point property holds when considering possibly open fixed points. The problem of counting fixed points in the closed setting remains open, but we provide sufficient conditions for a λ -term to have either one or infinitely many fixed points. In the main result of this paper we prove that in every sensible λ -theory there exists a λ -term that violates the fixed point property.

We then study the open problem concerning the existence of a double fixed point combinator and propose a proof technique that could lead towards a negative solution. We consider interpretations of the λY -calculus into the λ -calculus together with two reduction extension properties, whose validity would entail the non-existence of any double fixed point combinators. We conjecture that both properties hold when typed λY -terms are interpreted by arbitrary fixed point combinators. We prove reduction extension property I for a large class of fixed point combinators.

Finally, we prove that the λY -theory generated by the equation characterizing double fixed point combinators is a conservative extension of the λ -calculus.

1 Introduction

A fundamental result in the λ -calculus is the fixed point theorem [2, Thm. 2.1.5] stating that every λ -term M has at least one *fixed point*, i.e. a λ -term X satisfying $MX =_{\beta} X$. The λ -calculus also enjoys the *range property* [2, Thm. 20.2.5] stating that the range of every combinator (closed λ -term) M is either a singleton, when M represents a constant function, or infinite, in the sense that it contains denumerably many pairwise β -distinct λ -terms. It is therefore natural to wonder whether a similar property, which we call here ‘the fixed point property’, is enjoyed by the set of fixed points of an arbitrary closed λ -term:

Does every combinator have either one or infinitely many (closed) fixed points?

The above question appears as Problem 25 in the Typed Lambda Calculi and Applications (TLCA) list of open problems [16] and was first raised by Intrigila and Biasone in [17]; the first part of the present paper reports progress on this question. We first prove that if one considers *open* λ -terms,

then the question has a positive answer (Theorem 4.6). This result is not particularly difficult to achieve, but we believe it is interesting since it motivates the restriction to combinators and closed fixed points. For the more difficult question of closed fixed points, in [17] the authors prove that the fixed point property is satisfied by all combinators having a fixed point that is β -normalizable. We present several results in the same spirit. For example, we prove that the set of fixed points of a closed zero¹ λ -term is always infinite (Proposition 5.4) and that if a combinator has a fixed point that is a recurrent² zero λ -term then it has either one or infinitely many fixed points (Theorem 5.7).

The problem of determining whether the fixed point property or the range property holds radically changes when considering as equality between λ -terms an arbitrary λ -theory \mathcal{T} , i.e. an arbitrary context-closed extension of β -convertibility. Indeed, a set containing infinitely many β -distinct λ -terms might become finite modulo \mathcal{T} . For instance, it is well known that the range property is valid in every recursively enumerable λ -theory [2, Thm. 20.2.5] and in every λ -theory equating all λ -terms having the same Böhm tree (BT) [2, Thm. 20.2.6], while Polonsky recently proved that it fails in the λ -theory \mathcal{H} generated by equating all unsolvables [26]. This last result led Intrigila and Statman to conjecture in [18] that in the λ -theory \mathcal{H} ‘a very complicated example could exist with, say, exactly two fixed points’. In Corollary 6.3 we show that a λ -term satisfying such a property exists in every sensible λ -theory \mathcal{T} (in particular, in \mathcal{H}), thus proving their conjecture. Starting from this example, we are able to construct for every natural number $k > 0$ a λ -term having exactly k pairwise \mathcal{T} -distinct fixed points (Proposition 6.6). In [18], the authors also managed to construct in an ingenious, but complex way, a λ -theory satisfying the range property but not satisfying the fixed point property. An easy consequence of our result (Corollary 6.5) is that the same holds for the much more natural λ -theory \mathcal{B} generated by equating all λ -terms having the same BT, as it is obviously sensible and satisfies the range property by [2, Thm. 20.2.6].

The fixed point theorem of λ -calculus is a consequence of the existence of *fixed point combinators* that are λ -terms Y satisfying $YX =_{\beta} X(YX)$ for all λ -terms X . Clearly, every fixed point combinator Y satisfies the equation $\delta Y =_{\beta} Y$ where δ is the λ -term $\mathbf{SI} =_{\beta} \lambda yx.x(yx)$. Moreover, Böhm noticed that if Y is a fixed point combinator then $Y\delta$ is also. This consideration led Statman to raise in [29] the question of whether there exists a *double* fixed point combinator, namely a fixed point combinator Y satisfying $Y\delta =_{\beta} Y$. Intuitively, the application of δ has the effect of ‘slowing down’ the head reduction of Y and this should entail that $Y\delta$ and Y cannot have a common reduct. For this reason Statman conjectured that double fixed point combinators do not exist. A proof of Statman’s conjecture has been suggested by Intrigila in [15]. However, in 2011, Endrullis [8] has discovered a gap in a crucial case of the argument. The problem is therefore considered open.

The second part of the paper is devoted to presenting a proof technique that we believe will be useful in settling Statman’s conjecture. The main technical tool that we use is the λY -calculus [1, Section 6.1], a classic extension of the λ -calculus with a unary constant Y behaving as a fixed point combinator. We first show that the λY -calculus can be soundly interpreted in the λ -calculus, by replacing a fixed point operator for each occurrence of Y in a λY -term M . We then define two properties of such an interpretation map, which we call ‘reduction extension properties’, and we analyse under what circumstances they actually hold. On the one hand, we are able to prove that property I holds for a large class of reducing fixed point combinators (Corollary 7.31), including all putative double fixed point combinators. On the other hand, it is not difficult to check that property II fails in the untyped setting because the interpretation map is not injective. We conjecture

¹Intuitively, *zero λ -terms* are λ -terms that cannot be converted to an abstraction. We refer to Section 5.2 for a more thorough discussion about this terminology.

²A λ -term M is *recurrent* if, for all λ -terms $N, M \rightarrow_{\beta} N$ entails $N \rightarrow_{\beta} M$ (this notion is due to M. Venturini-Zilli).

however that a generalized version of both properties (Definition 7.27) holds for all fixed point combinators in the simply typed setting, and we show that this would entail the non-existence of double fixed point combinators (as discussed at the end of Section 7).

Finally, we analyse the question of whether the λY -theory δ^* generated by the equation $Yx = Y\delta x$ (the equation characterizing double fixed point combinators) is a conservative extension of the λ -calculus. Indeed, as discussed in Section 8, a negative answer would entail the non-existence of double fixed point combinators. Unfortunately, it turns out that the answer is positive, as shown in Theorem 8.9.

2 Preliminaries

In this preliminary section we introduce some notions and notations that are used in the rest of the article.

2.1 Lambda calculus

For the λ -calculus we mainly use the notations of Barendregt's first book [2].

Let us fix an infinite set Var of variables. The set Λ of λ -terms is generated by

$$\Lambda : \quad M, N := x \mid \lambda x.M \mid MN \quad (\text{for } x \in \text{Var}).$$

As usual we assume that application associates to the left and has a higher precedence than λ -abstraction. For instance, we write $\lambda xyz.xyz$ for $\lambda x.(\lambda y.(\lambda z.((xy)z)))$.

NOTATION 2.1

We write $M^n N$ for $M(\dots(MN)\dots)$ and $NM^{\sim n}$ for $(\dots(NM)\dots)M$ (n times). In particular, for $n = 0$, we have $M^0 N = N = NM^{\sim 0}$.

The set $\text{FV}(M)$ of *free variables* of M and α -conversion are defined as in [2, Section 2.1]. We say that a λ -term M is *closed* whenever $\text{FV}(M) = \emptyset$ and we denote by Λ^o the set of all closed λ -terms. The set of *positions*, denoted $\text{pos}(M)$, in a λ -term M is the subset of $\{0, 1\}^*$ defined inductively by $\text{pos}(x) = \{\epsilon\}$, $\text{pos}(\lambda x.M) = \{\epsilon\} \cup 0 \cdot \text{pos}(M)$ and $\text{pos}(MN) = \{\epsilon\} \cup 0 \cdot \text{pos}(M) \cup 1 \cdot \text{pos}(N)$. If M is a λ -term and p is a position in M , the *subterm of M at p* is defined in the obvious way.

CONVENTION

Hereafter, we consider λ -terms up to α -conversion and we adopt Barendregt's variable convention [2, Conv. 2.1.13].

By historical tradition, any binary relation on Λ is called a *notion of reduction* on Λ . We say that a notion of reduction $r \subseteq \Lambda \times \Lambda$ is *compatible* (or *contextual*) whenever it is compatible with respect to the operations of application and lambda abstraction. A *reduction relation* on Λ is any compatible notion of reduction.

The main compatible relation of the λ -calculus is the β -relation \rightarrow_β , which is the compatible closure of the following notion of reduction:

$$(\lambda x.M)N \rightarrow M[N/x], \quad (\beta)$$

where $M[N/x]$ denotes the λ -term obtained by simultaneously substituting all free occurrences of x in M for N , subject to the usual proviso of avoiding capture of free variables in N . The

η -relation \rightarrow_η is the compatible closure of

$$\lambda x.Mx \rightarrow M \quad (\text{for } x \notin \text{FV}(M)). \quad (\eta)$$

Concerning specific combinators we fix the following notations:

$$\begin{aligned} \mathbf{I} &= \lambda x.x, & \mathbf{K} &= \lambda xy.x, & \mathbf{F} &= \lambda xy.y, & \mathbf{B} &= \lambda fgx.f(g(x)), & \mathbf{S} &= \lambda xyz.xz(yz), \\ \mathbf{\Delta} &= \lambda x.xx, & \mathbf{\Omega} &= \mathbf{\Delta}\mathbf{\Delta}, & \mathbf{\Delta}_3 &= \lambda x.xx, & \mathbf{\Omega}_3 &= \mathbf{\Delta}_3\mathbf{\Delta}_3, & \delta &= \lambda yx.x(yx), \end{aligned}$$

where \mathbf{I} is the identity, \mathbf{K} and \mathbf{S} are the combinators of combinatory logic, \mathbf{F} is the second projection, \mathbf{B} the functional composition, $\mathbf{\Omega}$ the paradigmatic looping λ -term and $\mathbf{\Omega}_3$ the ‘garbage’ producing looping λ -term. It is easy to check that δ is the β -normal form of \mathbf{SI} . We denote the n -th Church numeral by \mathbf{c}_n [2, Def. 6.4.4]. The symbol $=$ denotes definitional equality (possibly modulo α -conversion).

The pairing is encoded in the λ -calculus as follows (for $x \notin \text{FV}(MN)$):

$$[M, N] = \lambda x.xMN, \quad \text{with projections } \pi_1 = \lambda x.x\mathbf{K} \text{ and } \pi_2 = \lambda x.x\mathbf{F}.$$

For instance, $\pi_1[M_1, M_2] \rightarrow_\beta [M_1, M_2]\mathbf{K} \rightarrow_\beta \mathbf{K}M_1M_2 \rightarrow_\beta (\lambda y.M_1)M_2 \rightarrow_\beta M_1$.

2.2 Rewriting

Given a reduction relation \rightarrow_r , we denote its transitive and reflexive closure by \twoheadrightarrow_r and its transitive, symmetric and reflexive closure by $=_r$. The relation \twoheadrightarrow_r is called *multistep r -reduction*, while $=_r$ is called *r -conversion*. We write $_r\leftarrow$ (resp. $_r\leftarrow$) for the relational inverse of \rightarrow_r (resp. \twoheadrightarrow_r) and \leftrightarrow_r for the symmetric closure of \rightarrow_r , i.e. $\rightarrow_r \cup {}_r\leftarrow$. Given two reduction relations \rightarrow_r and $\rightarrow_{r'}$, we write $\rightarrow_{rr'}$ for the relation $\rightarrow_r \cup \rightarrow_{r'}$. Similarly, we denote by $=_{rr'}$ the least contextual relation including $=_r \cup =_{r'}$.

DEFINITION 2.2

We recall the following standard auxiliary definitions.

- Given a notion of reduction \rightarrow , a *redex* is any term R such that $R \rightarrow P$ for some term P . For any term M , a *redex in M* is a pair $(C[], R)$ where $C[]$ is a one-hole context such that $M = C[R]$ and R is a redex.
- Given a reduction relation \rightarrow_r and two terms M and N such that $M \twoheadrightarrow_r N$, we call any witness $M = M_0 \rightarrow_r M_1 \rightarrow_r \dots \rightarrow_r M_n = N$ of $M \twoheadrightarrow_r N$ a *reduction sequence* from M to N . *Par abus de langage*, we shall occasionally refer to $M \twoheadrightarrow_r N$ as a reduction sequence without specifying the witness.
- Given a term M and a reduction relation \rightarrow_r , the *reduction graph* of M , denoted $\mathcal{G}_r(M)$ is the directed graph whose nodes are all terms N such that $M \twoheadrightarrow_r N$ and there is an edge from node P to node Q if $P \rightarrow_r Q$.
- A finite or infinite sequence

$$M = M_0 \rightarrow_r M_1 \rightarrow_r M_2 \rightarrow_r \dots$$

is called *cofinal* in $\mathcal{G}_r(M)$ if, for every node P of $\mathcal{G}_r(M)$, there is a directed path in $\mathcal{G}_r(M)$ from P to some M_i .

- As usual, for a step $M \rightarrow_\beta N$, the *residual relation* maps every set \mathcal{F} of β -redexes in M to a set of β -redexes in N , the set of *residuals of \mathcal{F} across the step*;³ the relation extends transitively to reduction sequences $M \rightarrow_\beta N$ in the obvious way.
- A *development* of a set of redexes \mathcal{F} in M is a reduction sequence $M \rightarrow_\beta M_1 \rightarrow_\beta M_2 \rightarrow_\beta \dots$ such that every step in the sequence is the contraction of a residual of a redex in \mathcal{F} .
- A development of a set of redexes \mathcal{F} in a term M is *complete* if it is finite and its final term has an empty set of residuals of \mathcal{F} across the sequence. By standard results, all maximal developments of \mathcal{F} are complete, hence finite, and all complete developments of \mathcal{F} end in the same term. Furthermore, if \mathcal{F} and \mathcal{G} are sets of redexes in a term M , the set of residuals of \mathcal{G} is the same across any complete development of \mathcal{F} , and is denoted \mathcal{G}/\mathcal{F} .
- A reduction sequence $M = M_0 \rightarrow_\beta M_1 \rightarrow_\beta M_2 \rightarrow_\beta \dots$ is *standard* if, for all i, j with $j < i$, the redex contracted in the step $M_i \rightarrow_\beta M_{i+1}$ is not a residual across $M_j \rightarrow_\beta \dots \rightarrow_\beta M_i$ of any redex to the left (in M_j) of the redex contracted in $M_j \rightarrow_\beta M_{j+1}$ (i.e. intuitively in a standard reduction, leftmost-outermost redexes are contracted first).
- *Permutation equivalence* is the smallest equivalence relation \equiv on reduction sequences such that
 1. $\rho; \sigma; \tau \equiv \rho; \sigma'; \tau$ whenever $\sigma \equiv \sigma'$ and
 2. if \mathcal{F} and \mathcal{G} are sets of redexes of the same term, then $\sigma \equiv \tau$ whenever σ is obtained by first performing a complete development of \mathcal{F} followed by a complete development of \mathcal{G}/\mathcal{F} , and τ is obtained by first performing a complete development of \mathcal{G} followed by a complete development of \mathcal{F}/\mathcal{G} .
- A *redex with history* is a pair $(M \rightarrow_\beta N, R)$ consisting of a reduction sequence $M \rightarrow_\beta N$ and a redex R in N . A redex with history $(M \rightarrow_\beta P, S)$ is a *copy* of a redex with history $(M \rightarrow_\beta N, R)$ if there is a reduction sequence $N \rightarrow_\beta P$ such that (i) $M \rightarrow_\beta N \rightarrow_\beta P$ is permutation equivalent to $M \rightarrow_\beta P$, and (ii) S is a residual of R across $N \rightarrow_\beta P$. The symmetric and transitive closure of the copy relation is called the *family relation* on redexes with history and is obviously an equivalence relation. If two redexes with history are elements of the same equivalence class in the family relation they are said, *par abus de langage*, to *belong to the same family relation*.

REMARK 2.3

It is easy to check that $M =_r N$ if and only if there exists a sequence $M = M_0 \leftrightarrow_r M_1 \leftrightarrow_r \dots \leftrightarrow_r M_k = N$ of length $k \geq 0$.

2.3 Solvability

Lambda terms are classified as solvable or unsolvable, depending on their capability of interaction with the environment.

DEFINITION 2.4

A closed λ -term M is *solvable* if there are $P_1, \dots, P_k \in \Lambda$ such that $MP_1 \dots P_k =_\beta \mathbf{I}$. An open λ -term M is *solvable* if its closure $\lambda x_1 \dots x_n. M$ is.

We say that a λ -term M is in *head normal form (hnf)* if it has the shape $\lambda x_1 \dots x_n. x_i. M_1 \dots M_k$ where $n, k \geq 0$ and either $1 \leq i \leq n$ or x_i occurs freely. We say that M *has an hnf* whenever

³We omit the details, see [2, Ch. 11.2].

$M \rightarrow_\beta N$ for some N in hnf. It is well known that if a λ -term has an hnf, then such an hnf can be obtained by repeatedly reducing its *head redex* $\lambda x_1 \dots x_n. (\lambda x. M) N M_1 \dots M_k$. Solvability has been characterized in terms of head normalization by Wadsworth.

THEOREM 2.5 (Wadsworth [31]).

A λ -term M is solvable if and only if it has a hnf.

Every closed λ -term M can be turned into an unsolvable one by applying enough Ω 's. In other words, for k large enough, $M\Omega^{\sim k}$ is unsolvable [2, Lemma 17.4.4]. The following lemma will be useful in Section 6 and is a revisitation of such a result.

LEMMA 2.6

Let $M \in \Lambda$ and $y \in \text{Var}$. If $My\Omega^{\sim n}$ is solvable for all $n \in \mathbb{N}$, then $M =_\beta \lambda x_0 \dots x_k. x' M_1 \dots M_m$ for some $k, m \geq 0$ and $x' \in \text{FV}(M) \cup \{x_0\}$.

PROOF. For $n = 0$ we have that My is solvable, which entails that M has an hnf $\lambda x_0 \dots x_k. x' M_1 \dots M_m$. Toward contradiction, suppose $x' = x_j$, with $0 < j \leq k$. Then for the appropriate $M'_1, \dots, M'_m \in \Lambda$ we have $My\Omega^{\sim k} =_\beta \Omega M'_1 \dots M'_m$, which is unsolvable. This contradicts the hypothesis for $n = k$. \square

2.4 Lambda theories

The equational theories of the untyped λ -calculus are called λ -theories and become the main object of study when considering the equivalence between λ -terms more important than the process of computation.

More precisely, we will be considering *congruences*, which are compatible binary equivalence relations on Λ .

DEFINITION 2.7

A λ -theory \mathcal{T} is any congruence on Λ containing the β -conversion.

As a matter of notation, we write $\mathcal{T} \vdash M = N$ or just $M =_{\mathcal{T}} N$ for $(M, N) \in \mathcal{T}$. Let \mathcal{T} be a λ -theory and M be a λ -term, we write $\Lambda_{\mathcal{T}}$ for the set Λ modulo \mathcal{T} and $[M]_{\mathcal{T}}$ for the \mathcal{T} -equivalence class of M . Similarly, we set $\Lambda_{\mathcal{T}}^o = \{[M]_{\mathcal{T}} \mid M \in \Lambda^o\}$. Given a subset $X \subseteq \Lambda_{\mathcal{T}}$, we write $M \in_{\mathcal{T}} X$ whenever $[M]_{\mathcal{T}} \in X$.

The set of all λ -theories, ordered by set-theoretical inclusion, constitutes a complete lattice $\lambda\mathcal{T}$ of cardinality 2^{\aleph_0} . As shown by Salibra and his coauthors in their works [22, 23, 27], $\lambda\mathcal{T}$ has a very rich mathematical structure. The lattice $\lambda\mathcal{T}$ has a bottom element $\lambda\beta$ that equates only β -convertible λ -terms, and a top element ∇ that equates all λ -terms.

DEFINITION 2.8

A λ -theory \mathcal{T} is

- *consistent* if $\mathcal{T} \neq \nabla$,
- *inconsistent* if it is not consistent,
- *sensible* if it equates all unsolvable terms,
- *extensional* whenever, for all λ -terms M, N and any variable $x \notin \text{FV}(MN)$, $Mx =_{\mathcal{T}} Nx$ implies $M =_{\mathcal{T}} N$.

CONVENTION

We will only consider consistent λ -theories and omit the assumption.

By [2, Thm. 2.1.29], \mathcal{T} is extensional exactly when it contains the η -conversion.

We denote by $\lambda\beta\eta$ the smallest extensional λ -theory and by \mathcal{H} the smallest sensible λ -theory. We denote by \mathcal{B} the λ -theory equating two λ -terms if and only if they have the same BT [2, Def. 10.1.4]. It is well known that \mathcal{H} also admits a unique maximal extension, which is denoted by \mathcal{H}^* [31]. As shown in [2, Thm. 17.4.16], the strict inclusions $\mathcal{H} \subsetneq \mathcal{B} \subsetneq \mathcal{H}^*$ hold.

The λ -theories \mathcal{H} , \mathcal{B} and \mathcal{H}^* have been extensively studied in the literature. In particular, Hyland proved in [14] that two λ -terms M and N are equal in \mathcal{H}^* exactly when their BTs are equal up to ‘possibly infinite’ η -expansions (see also [2, Thm. 16.2.7]). As an easy consequence, we get the following remark that will be used in Section 6.

REMARK 2.9

Let \mathcal{T} be a sensible λ -theory. For all $M, N \in \Lambda$, if $\mathcal{T} \vdash M = N$ then one of the following conditions holds:

- (i) $M =_{\mathcal{T}} N =_{\mathcal{T}} \Omega$,
- (ii) There are $k, m \geq 0$ such that

$$M =_{\beta\eta} \lambda x_1 \dots x_k . y M_1 \dots M_m \quad \text{and} \quad N =_{\beta\eta} \lambda x_1 \dots x_k . y N_1 \dots N_m,$$

where $\mathcal{T} \vdash M_i = N_i$ for all $1 \leq i \leq m$.

By condition (ii), if $M =_{\mathcal{T}} \lambda x_1 \dots x_{k_1} . y M_1 \dots M_{m_1}$ and $N =_{\mathcal{T}} \lambda x_1 \dots x_{k_2} . y N_1 \dots N_{m_2}$ then $m_1 - k_1 = m_2 - k_2$. Intuitively, this means that the number of λ -abstractions and applications can be matched by performing some η -expansions.

3 Fixed points and fixed point combinators

In λ -calculus a fixed point of a λ -term F is an $X \in \Lambda$ satisfying $FX =_{\beta} X$. The fixed point theorem states that all λ -terms have a fixed point [2, Thm. 2.1.5], a result that follows from the existence of fixed point combinators.

THEOREM 3.1

For every λ -term M , there exists X such that $MX =_{\beta} X$. Actually, there exists a closed λ -term Y such that for any λ -term M , $M(YM) =_{\beta} YM$.

In this section we start by defining fixed points relative to some λ -theory \mathcal{T} , and then provide some notions of fixed point combinators and examples.

DEFINITION 3.2

Let \mathcal{T} be a λ -theory.

- (1) Given two λ -terms M, N , we say that N is a *fixed point of M in \mathcal{T}* whenever $MN =_{\mathcal{T}} N$.
- (2) For $M \in \Lambda$, we let $\text{Fix}_{\mathcal{T}}(M) = \{[N]_{\mathcal{T}} \mid N \in \Lambda, MN =_{\mathcal{T}} N\}$ be the set of all (\mathcal{T} -classes of) fixed points of M in \mathcal{T} .
- (3) Similarly, for $M \in \Lambda^o$, we let $\text{Fix}_{\mathcal{T}}^o(M) = \text{Fix}_{\mathcal{T}}(M) \cap \Lambda_{\mathcal{T}}^o$ be the set of (\mathcal{T} -classes of) all closed fixed points of M in \mathcal{T} .

When $\mathcal{T} = \lambda\beta$ we simply say that N is a fixed point of M and write $\text{Fix}(M)$ and $\text{Fix}^o(M)$ for the set of its open and closed fixed points, respectively.

REMARK 3.3

Given a λ -theory \mathcal{T} and λ -terms M, N , if $N \in_{\mathcal{T}} \text{Fix}_{\mathcal{T}}(M)$ then for all λ -theories $\mathcal{T}' \supseteq \mathcal{T}$ we have $N \in_{\mathcal{T}'} \text{Fix}_{\mathcal{T}'}(M)$. In particular, if N is a fixed point of M we have $N \in_{\mathcal{T}} \text{Fix}_{\mathcal{T}}(M)$ for all λ -theories \mathcal{T} .

EXAMPLE 3.4

- (i) Since $\mathbf{I}M =_{\beta} M$ for all $M \in \Lambda$, we have that every λ -term is a fixed point of the identity \mathbf{I} . Therefore, $\text{Fix}(\mathbf{I}) = \Lambda_{\lambda\beta}$ and $\text{Fix}^o(\mathbf{I}) = \Lambda_{\lambda\beta}^o$.
- (ii) Since $\mathbf{F}M =_{\beta} \mathbf{I}$ for all $M \in \Lambda$, we have that only λ -terms β -convertible with \mathbf{I} are fixed points of \mathbf{F} and therefore that both $\text{Fix}(\mathbf{F})$ and $\text{Fix}^o(\mathbf{F})$ are singletons.

3.1 Fixed point combinators

As shown in the fixed point theorem, every λ -term has at least one fixed point, since fixed points can be constructed through fixed point combinators.

DEFINITION 3.5

- (i) A λ -term Y is a *fixed point combinator* (or *fpc*) if $Yx =_{\beta} x(Yx)$ for every $x \notin \text{FV}(Y)$;
- (ii) An fpc Y is *reducing* if $Yx \rightarrow_{\beta} x(Yx)$ for every $x \notin \text{FV}(Y)$;
- (iii) An fpc Y is *terminal* if it is reducing and there is a reduction $\rho : Yx \rightarrow_{\beta} x(Yx)$ with the property that the sequence of terms in the infinite reduction

$$Yx \xrightarrow{\rho}_{\beta} x(Yx) \xrightarrow{x\rho}_{\beta} x^2(Y\rho) \xrightarrow{x(x\rho)}_{\beta} x^3(Yx) \xrightarrow{x^3(\rho)}_{\beta} \dots$$

is cofinal in the reduction graph $\mathcal{G}_{\beta}(Yx)$.

Note that, following a well-established tradition [9, 10], we do not require that fpc's are actual combinators in the sense of being closed λ -terms. From the existence of closed fpc's Y it follows however that $YM \in_{\lambda\beta} \text{Fix}(M)$; therefore, $\text{Fix}_{\mathcal{T}}(M) \neq \emptyset$ (resp. $\text{Fix}_{\mathcal{T}}^o(M) \neq \emptyset$) for all (closed) λ -terms M .

DEFINITION 3.6

Let $\mathcal{T} \in \lambda\mathcal{T}$ and $M \in \Lambda$. A fixed point $N \in_{\mathcal{T}} \text{Fix}_{\mathcal{T}}(M)$ is called *canonical* if $N =_{\mathcal{T}} YM$ for some fpc Y .

We now provide some examples of open and closed fpc's, reducing and non-reducing fpc's and terminal and non-terminal fpc's.

EXAMPLE 3.7

- Curry's fixed point combinator $\mathbf{Y} = \lambda f. \Delta_f \Delta_f$ where $\Delta_f = \lambda x. f(xx)$, which is closed and not reducing.
- Geuvers and Verkoelen's fixed point combinator $\lambda f. (\Delta(\lambda xy. f(yxy)))\Delta$ defined in [12] is also closed and not reducing.
- Turing's fixed point combinator $\Theta = \mathbf{W}\mathbf{W}$ where $\mathbf{W} = \lambda wx. x(wx)$, which is closed and reducing.
- Turing's fpc can be parametrized by setting $\Theta_M = \mathbf{V}\mathbf{V}M$ for $M \in \Lambda$ and $\mathbf{V} = \lambda vpx. x(vvp x)$. Indeed $\Theta_M x = \mathbf{V}\mathbf{V}Mx \rightarrow_{\beta} x(\mathbf{V}\mathbf{V}Mx) = x(\Theta_M x)$, so Θ_M is a reducing fpc for all $M \in \Lambda$. Notice that for any variable z , Θ_z is open and terminal, while Θ_{Ω_3} is closed and not terminal.

- Polonsky's fpc is introduced here and works for arbitrary $A, B \in \Lambda$. The fpc is the λ -term XZ where (recall that $[M, N] = \lambda z.zMN$ for $z \notin \text{FV}(MN)$):

$$X = \lambda yx.x(y(yA\mathbf{F})\mathbf{K}(\lambda z.[z, yB\mathbf{F}])x) \text{ and } Z = \lambda x.[x, X].$$

Note that $ZM\mathbf{K} \rightarrow_\beta M$ and $ZM\mathbf{F} \rightarrow_\beta X$ hold. The fpc XZ is reducing:

$$XZx \rightarrow_\beta x(Z(ZA\mathbf{F})\mathbf{K}(\lambda z.[z, ZB\mathbf{F}])x) \rightarrow_\beta x(ZX\mathbf{K}(\lambda z.[z, X])x) \rightarrow_\beta x(XZx).$$

Whether XZ is closed or terminal depends on the chosen $A, B \in \Lambda$.

It is easy to check that all fpc's have the same BT; therefore, all canonical fixed points are equated in every λ -theory $\mathcal{T} \supseteq \mathcal{B}$. There are however λ -terms that are not fpc's but have the same BT as a fixed point combinator; such terms are called weak fixed point combinators (or *looping combinators* in [7, 13]):

DEFINITION 3.8

A λ -term Y is a *weak fixed point combinator* if, for all $x \notin \text{FV}(Y)$:

$$Yx =_{\mathcal{B}} x(Yx).$$

Since the BT of a weak fpc is equal to that of an fpc, the following alternative characterization of weak fpc's is easily obtained.

PROPOSITION 3.9

A λ -term Y is a weak fpc if and only if there exists a family of λ -terms $(Y_i)_{i \in \mathbb{N}}$ such that $Y = Y_0$ and, for all $i \in \mathbb{N}$ and x fresh, $Y_i x =_\beta x(Y_{i+1}x)$.

PROOF. (\Leftarrow) is trivial while (\Rightarrow) is an easy coinductive argument. \square

Since all the λ -terms Y_i 's above are weak fpc's themselves, this gives us the following coinductive characterization of weak fixed point combinators: a λ -term Y is a weak fpc if and only if $Yx =_\beta x(Y'x)$ for some weak fpc Y' and $x \notin \text{FV}(Y)$.

EXAMPLE 3.10

Define by double recursion [19], two β -distinct λ -terms Y and Y' such that $Yx =_\beta x(Y'x)$ and $Y'x =_\beta x(Yx)$. Then, both Y and Y' are weak fpc's.

Dealing with fpc's and weak fpc's suggests the following notions.

DEFINITION 3.11

Let $M \in \Lambda$.

- A variable $x \in \text{FV}(M)$ *eventually disappears from* M , written $x \notin_\beta M$ if there exists M' such that $M \rightarrow_\beta M'$ and $x \notin \text{FV}(M')$.
- Given $k \in \mathbb{N}$, we say that M is *k-constant* if $x \notin_\beta M^k(x)$, for $x \notin \text{FV}(M)$.

Clearly, if M is k -constant for some $k \in \mathbb{N}$ and $x \notin \text{FV}(M)$ then $x \notin_\beta \Theta_x M$. By exploiting this fact, we prove in Corollary 4.4 that for every k -constant λ -term M the set $\text{Fix}(M)$ is a singleton, thus generalizing Example 3.4(ii).

3.2 Derived fixed point combinators

An interesting line of research [19], consists in defining new fixed point combinators starting from existing ones. Notice, for instance, that $\Delta_\delta = \lambda w.\delta(w w) =_\beta \mathbf{W}$ where $\delta = \lambda yx.x(yx)$; therefore, $\mathbf{Y}\delta =_\beta (\lambda x.\delta(xx))(\lambda x.\delta(xx)) =_\beta \Theta$. In other words, Turing's fixed point combinator can be obtained from Curry's one by applying δ .

The following properties concerning the interaction between fpc's and δ have been pointed out by Böhm (see [2, Lemma 6.5.3]).

LEMMA 3.12

Let $Y \in \Lambda$.

- (i) Y is an fpc if and only if $\delta Y =_\beta Y$,
- (ii) if Y is a (reducing) fpc then also $Y\delta$ is.

Statman raised in [29] the following natural question and conjectured that it has a negative answer. (This question will be discussed more thoroughly in Section 7.)

PROBLEM 1

Is there a *double fpc*, which is an fpc Y satisfying $Y =_\beta Y\delta$?

This problem is interesting because Lemma 3.12 tells us that starting from an fpc Y ; it is always possible to define infinitely many fpc's $(Y_n)_{n \in \mathbb{N}}$ by setting

$$Y_0 = Y, \quad Y_{n+1} = Y_n\delta.$$

The difficult part is to prove that all the fpc's so obtained are β -distinct, a result that would clearly follow from Statman's conjecture. In the following case we know the answer, but the general case is an open question.

EXAMPLE 3.13

The *Scott sequence* $(Y_n)_{n \in \mathbb{N}}$ is generated by taking as Y_0 Curry's fpc \mathbf{Y} . As mentioned earlier, Turing's fpc Θ occurs as Y_1 in such a sequence. As shown by Klop in [19, Thm. 2.1] with an *ad hoc* argument, the Scott sequence contains no repetitions (i.e. $Y_i =_\beta Y_j$ if and only if $i = j$).

Other fpc's can be found starting from existing ones by mechanical search.

EXAMPLE 3.14

Let $Y \in \Lambda$ be an fpc. Klop's *Bible*⁴ fixed point combinator is given by $\mathbf{B} = \lambda e.\mathbf{B}Y\mathbf{B}eL$, where \mathbf{B} is the composition, and works for arbitrary $L \in \Lambda$. Notice that L remains in passive position during the reduction:

$$\mathbf{B}x \rightarrow_\beta \mathbf{B}Y\mathbf{B}xL \rightarrow_\beta Y(\mathbf{B}x)L =_\beta \mathbf{B}x(Y(\mathbf{B}x))L \rightarrow_\beta x(Y(\mathbf{B}x)L) =_\beta x(\mathbf{B}x).$$

3.3 The fixed point property

We have seen in Example 3.4 that, on the one hand, there are λ -terms having infinitely many fixed points, like the identity \mathbf{I} . On the other hand, there are λ -terms M possessing only one fixed point, namely those having a constant output like the second projection \mathbf{F} . Indeed, whenever there is an

⁴The name of such a combinator comes from the Dutch translation of 'bible', namely *bijbel*.

M' such that $MN =_{\beta} M'$ for all N , we have that $\text{Fix}(M)$ is a singleton. Therefore, it makes sense to wonder how many fixed points a λ -term M possesses, as Intrigila and Biasone did (in the closed case) [17].

The following terminology is inspired by the range property of λ -calculus [2, Thm. 17.1.16].

DEFINITION 3.15

Let \mathcal{T} be a λ -theory.

- A closed λ -term M has the *fixed point property* (*fpp*) in \mathcal{T} whenever $\text{Fix}_{\mathcal{T}}^o(M)$ is either a singleton or infinite.
- A λ -term M has the *open fixed point property* in \mathcal{T} if $\text{Fix}_{\mathcal{T}}(M)$ is either a singleton or infinite.
- The λ -theory \mathcal{T} satisfies the *fixed point property* (resp. *open fpp*) if every closed λ -term (resp. possibly open λ -term) has the same property in \mathcal{T} .

As usual, when \mathcal{T} is omitted, we assume that we are considering $\mathcal{T} = \lambda\beta$. In this terminology, the Problem 25 of the TLCA list can be rephrased as follows.

PROBLEM 2 ([16]).

Does $\lambda\beta$ satisfy the fixed point property?

Some modest advances on this problem are presented in Section 5, while in the next section we give a positive answer to the analogue question concerning the *open* fixed point property. However, we will be also interested in the following generalization of Problem 2 to arbitrary λ -theories.

PROBLEM 3

What are the λ -theories satisfying the fixed point property?

In Section 6 we will show that no sensible λ -theory \mathcal{T} satisfies the fixed point property (Theorem 6.4).

4 Canonical open fixed points are not normal

In this section we show that $\lambda\beta$ satisfies the open fixed point property. More precisely, we show that every λ -term exhibiting a non-constant behaviour has infinitely many canonical fixed points. Such a result is not particularly difficult to prove and motivates the choice made by Intrigila and Biasone of raising the question for closed fixed points only. (cf. [18], where such a property is proved for a λ -calculus having infinitely many constants.)

The proof relies on the following property of Turing's parametrized fpc, which will have interesting consequences for closed fixed points as well (e.g. Proposition 5.4).

LEMMA 4.1

For all $M, N \in \Lambda$, we have $\Theta_M =_{\beta} \Theta_N$ if and only if $M =_{\beta} N$.

PROOF. (\Rightarrow) First, notice that the head reduction of Θ_{zx} is given by

$$\Theta_{zx} = \mathbf{VV}zx \rightarrow_{\beta} (\lambda zx.x(\mathbf{VV}zx))zx \rightarrow_{\beta} (\lambda x.x(\mathbf{VV}zx))x \rightarrow_{\beta} x(\mathbf{VV}zx). \quad (3)$$

Suppose now that $\Theta_M =_{\beta} \Theta_N$ holds, then there are two standard reductions ρ, σ from Θ_{Mx} and Θ_{Nx} toward a common reduct X , namely

$$\Theta_{Mx} = \mathbf{VVM}x \xrightarrow{\rho}_{\beta} X \xleftarrow{\sigma}_{\beta} \mathbf{VVN}x = \Theta_{Nx}.$$

Each of these reductions must again factor through an initial segment of (3) and there are two subcases. If this segment is empty, then ρ and σ are actually internal reductions. By inspection, the only subterms of Θ_{Mx} , Θ_{Nx} that may have redexes are M and N , respectively. Thus, ρ and σ yield a confluence between M and N , so we are done.

Otherwise, ρ and σ factor through a segment of (3) of the same length (in order to result in the same shape of the final λ -term). In this case, the internal reductions which follow the segment are again a confluence between Θ_{Mx} and Θ_{Nx} , allowing us to conclude by induction hypothesis.

(\Leftarrow) Trivial. \square

LEMMA 4.2

Let $M \in \Lambda$ and $z \notin \text{FV}(M)$. If $z \notin \beta \Theta_z M$ then $\text{Fix}(M)$ is a singleton.

PROOF. Let $\sigma : \Theta_z M \rightarrow_\beta X$ be a standard reduction, with $z \notin \text{FV}(X)$.

We consider the projection of σ across the canonical reduction sequence

$$\begin{aligned} \Theta_z M &\rightarrow_\beta M(\Theta_z M) \rightarrow_\beta M(M(\Theta_z)) \rightarrow_\beta \cdots \\ &\rightarrow_\beta M^k(\Theta_z M) = M^k(\mathbf{VV}z M) \rightarrow_\beta \cdots \end{aligned}$$

Notice that the redex \mathbf{VV} occurring inside each term in the sequence above is created during the contraction of this redex in the previous term.

In particular, for any given k , we know that such a redex could not have been contracted by any reduction starting with $\Theta_z M$ and shorter than k steps.

We now complete the projection diagram with $k = |\sigma|$, the length of σ :

$$\begin{array}{ccc} \Theta_z M & \twoheadrightarrow & M^{|\sigma|}(\underline{\mathbf{VV}z M}) \\ \sigma \downarrow & & \downarrow \rho \\ X & \twoheadrightarrow & Z \end{array}$$

As just observed, the underlined redex cannot stand in the family relation to any redex contracted in σ (since it requires $|\sigma|$ redex contractions to be created).

Therefore, this redex remains untouched by the reduction ρ . As a result, the reduction $\rho : M^{|\sigma|}(\mathbf{VV}z M) \rightarrow_\beta Z$ lifts as $(\rho_0; \cdots; \rho_\ell)[\mathbf{VV}z/v]$, where

$$\begin{aligned} \rho_0 : M^{|\sigma|}(vM) &\rightarrow_\beta Z_0[vM, \dots, vM], & \rho_i : M &\rightarrow_\beta M_i \quad (1 \leq i \leq \ell) \\ Z &= Z_0[\mathbf{VV}zM_1, \dots, \mathbf{VV}zM_\ell]. \end{aligned}$$

But since z is not a free variable of X , it cannot occur in Z either. That is, we must have $\ell = 0$, and therefore the reduction ρ_0 is of the form:

$$\rho_0 : M^{|\sigma|}(vM) \rightarrow_\beta Z_0, \quad \text{for } v \notin \text{FV}(Z_0).$$

Putting it all together, we find

$$\Theta_z M \rightarrow_\beta M^k(vM)[\mathbf{VV}z/v] \rightarrow_\beta Z_0[\mathbf{VV}z/v] = Z_0.$$

Now we are done, since for $X =_\beta MX$ and $k = |\sigma|$, we have

$$X =_\beta M^k(X) = M^k(vM)[\mathbf{K}X/v] =_\beta Z_0[\mathbf{K}X/v] = Z_0,$$

whence all fixed points of M are β -convertible with Z_0 . \square

PROPOSITION 4.3

Let $M \in \Lambda$ and let $y, z \notin \text{FV}(M)$ be distinct variables. If $\Theta_y M =_\beta \Theta_z M$ then the set $\text{Fix}(M)$ is a singleton.

PROOF. Since $\Theta_y M =_\beta \Theta_z M$ they have a common reduct X , i.e. $\Theta_y M \rightarrow_\beta X \leftarrow_\beta \Theta_z M$. Clearly neither y nor z can occur in X , so we conclude by Lemma 4.2. \square

Since $x \notin_\beta M^k(x)$ entails $z \notin_\beta \Theta_z M$, which in turn implies $\Theta_y M =_\beta \Theta_z M$, we obtain the following property of k -constant λ -terms.

COROLLARY 4.4

Let $k \in \mathbb{N}$. For every k -constant λ -term M the set $\text{Fix}(M)$ is a singleton.

THEOREM 4.5

For every λ -term M , either M is k -constant for some $k \in \mathbb{N}$, or it has infinitely many, pairwise distinct, canonical fixed points.

PROOF. If M is k -constant, then $\text{Fix}(M)$ is a singleton by Corollary 4.4. Otherwise, given a fresh variable x , every M' satisfying $\Theta_x M \rightarrow_\beta M'$ contains a free occurrence of x . This entails that $\Theta_y M \neq_\beta \Theta_z M$ for all distinct y, z that do not occur in M . Therefore, $\{\Theta_z M \mid z \in \text{Var} - \text{FV}(M)\} \subseteq \text{Fix}(M)$ and this set is infinite. \square

We obtain the following result concerning the open fpp for $\lambda\beta$.

THEOREM 4.6

The λ -theory $\lambda\beta$ satisfies the open fixed point property.

5 Some results concerning sets of fixed points in $\lambda\beta$

5.1 First observations

In this section we work in $\lambda\beta$.

EXAMPLE 5.1

The following examples are meant to illustrate the basic behaviour of the sets $\text{Fix}(M)$ and $\text{Fix}^o(M)$.

- (1) Define, for any $n \geq 1$, $\mathbf{App}_n = \lambda f x_1 \dots x_n. f x_1 \dots x_n$. Let $\mathfrak{o}_n \subseteq \Lambda$ be the set of λ -terms M such that $M =_\beta \lambda y_1 \dots y_n. N$ for some λ -term N .
(Notice that $\mathfrak{o}_{n+1} \subseteq \mathfrak{o}_n$ for each n ; the elements of $\mathfrak{o}_n - \mathfrak{o}_{n+1}$ are sometimes called *terms of order n* .)
If $M \in \mathfrak{o}_n$, then $M =_\beta \lambda y_1 \dots y_n. N$ for some N , and hence

$$\begin{aligned} \mathbf{App}_n M &=_\beta \lambda x_1 \dots x_n. M x_1 \dots x_n \\ &=_\beta \lambda x_1 \dots x_n. N[x_1/y_1, \dots, x_n/y_n] =_\alpha \lambda y_1 \dots y_n. N = M, \end{aligned}$$

whence $M \in_{\lambda\beta} \text{Fix}(\mathbf{App}_n)$.

Conversely, if $M \notin \mathfrak{o}_n$, then $\mathbf{App}_n M \rightarrow_\beta \lambda x_1 \dots x_n. M x_1 \dots x_n \in \mathfrak{o}_n$, and thus $M \notin_{\lambda\beta} \text{Fix}(\mathbf{App}_n)$. Hence, $\text{Fix}(\mathbf{App}_n) = \{[M]_\beta \mid M \in \mathfrak{o}_n\}$.

- (2) For all λ -terms F we prove that $\text{Fix}(F) \neq \{[\mathfrak{Q}]_\beta, [\mathbf{K}]_\beta\} \neq \text{Fix}^o(F)$.
Assume, by contradiction, that F satisfies $\text{Fix}(F) = \{[\mathfrak{Q}]_\beta, [\mathbf{K}]_\beta\}$. Observe that $[\mathfrak{Q}]_\beta = \{M \mid M \rightarrow_\beta \mathfrak{Q}\}$, whence $F \mathfrak{Q} =_\beta \mathfrak{Q}$ if and only if $F \mathfrak{Q} \rightarrow_\beta \mathfrak{Q}$. Split on cases according to the

solvability of F .

- If F is unsolvable then $F \rightarrow_{\beta} \lambda x. \mathbf{\Omega}$, but then $F \mathbf{K} =_{\beta} \mathbf{\Omega} \neq_{\beta} \mathbf{K}$, a contradiction.
 - If F is solvable, then $F \rightarrow_{\beta} \lambda x_1 \dots x_n. x_i N_1 \dots N_k$ for some $n, i, k \geq 0$. As $(\lambda x_1 \dots x_n. x_i N_1 \dots N_k) \mathbf{\Omega} \rightarrow_{\beta} \mathbf{\Omega}$, we must have $i, n = 1$ and $k = 0$. Hence, $F =_{\beta} \mathbf{I}$ and thus $\text{Fix}(F) = \Lambda_{\lambda\beta}$, a contradiction.
- (3) Let $x \notin \text{FV}(M)$ and $F = \lambda x. M$ (note that if M is closed, then so is F). Then if $N =_{\beta} M$ we have $F N =_{\beta} M =_{\beta} N$, hence $N \in_{\lambda\beta} \text{Fix}(F)$; and if $N \neq_{\beta} M$, then $F N = (\lambda x. M) N \rightarrow_{\beta} M[N/x] = M \neq_{\beta} N$, whence $N \notin_{\lambda\beta} \text{Fix}(F)$. Thus, for every $M \in \Lambda$, there exists a λ -term F such that $\text{Fix}(F) = \{[M]_{\beta}\}$. If $M \in \Lambda^o$, we may choose $F \in \Lambda^o$.
- (4) Define $\langle \cdot \rangle$ by $\langle T \rangle = \lambda z. zT$, where $z \notin \text{FV}(T)$. Set $F = \lambda x. x\langle x \rangle$, $X = \langle \mathbf{I} \rangle$ and $Z = \lambda y. y\langle y \rangle$. Then, we have

$$FX =_{\beta} X\langle X \rangle =_{\beta} \langle \mathbf{I} \rangle \langle \langle \mathbf{I} \rangle \rangle =_{\beta} \langle \langle \mathbf{I} \rangle \rangle \mathbf{I} =_{\beta} \mathbf{I} \langle \mathbf{I} \rangle =_{\beta} \langle \mathbf{I} \rangle = X$$

$$FZ =_{\beta} Z\langle Z \rangle =_{\beta} \lambda y. \langle Z \rangle \langle y \rangle =_{\beta} \lambda y. \langle y \rangle Z =_{\beta} \lambda y. Zy =_{\beta} Z$$

yet at the same time

$$X(\mathbf{KI})\mathbf{K} = \langle \mathbf{I} \rangle (\mathbf{KI})\mathbf{K} =_{\beta} (\mathbf{KI})\mathbf{IK} =_{\beta} \mathbf{IK} =_{\beta} \mathbf{K}$$

$$Z(\mathbf{KI})\mathbf{K} =_{\beta} \mathbf{KI} \langle \mathbf{K} \rangle =_{\beta} \mathbf{I}.$$

The last two equations show that $X \neq_{\beta} Z$. Hence, there is a closed λ -term F with $\text{Fix}_{\lambda\beta}(F) \neq \Lambda_{\beta}$ (since $F\mathbf{\Omega} \neq_{\beta} \mathbf{\Omega}$) such that there are at least two elements in $\text{Fix}_{\lambda\beta}^o(F)$ having distinct normal forms.

The first result of the section is that unless $\text{Fix}(F)$ and $\text{Fix}^o(F)$ are singletons, they cannot solely consist of equivalence classes of λ -terms in normal forms.

PROPOSITION 5.2

Let F be a closed λ -term. If $\text{Fix}(F)$ contains at least two elements, then at least one element does not have a normal form.

PROOF. By the fixed point theorem, F has at least one fixed point of the form YF for some fpc Y . We shall prove that if YF has a normal form, then F has at most one fixed point; the desideratum follows immediately from this.

Any λ -term having a normal form is an isolated point in the tree topology on Λ [2, Lem. 14.3.23]; hence, YF is isolated.

By the continuity theorem [2, Thm. 14.3.22], the map $X \mapsto XF$ is continuous, whence there is a neighbourhood of Y in the tree topology that is mapped to the singleton YF . As the BT of Y is $\lambda f. f(f(f(\dots)))$ and the tree topology has as basic opens all (extensions of) finite approximants of BTs (see, e.g. [2, Cor. 10.2.7]), there exists a $k > 0$ such that $(\lambda f. f^k(\mathbf{\Omega}))$ is mapped to YF . Hence, $(\lambda f. f^k(\mathbf{\Omega}))F = YF$ and $(\lambda f. f^k(\mathbf{\Omega}))F$ is a normalizing term.

By the genericity lemma [2, Prop. 14.3.24], there is a fresh variable z such that $F^k(\mathbf{\Omega}) = (\lambda f. f^k(\mathbf{\Omega}))F = F^k(z)$. As z is fresh, for any term M , we have $F^k(M) = (\lambda f. f^k(\mathbf{\Omega}))F$, and thus $F^k(M) = YF$.

If M is a fixed point of F , then $M =_{\beta} FM$, and thus $M =_{\beta} F^k(M) = YF$, concluding the proof. \square

The results in the rest of this section concern terms that have no weak hnf, namely terms that do not reduce to an abstraction regardless of which substitution is applied to them.

Meaningless Set	Unsolvable	\mathfrak{Z}	Mute
Definition	$\forall \vartheta : \text{Var} \rightarrow \Lambda, \vec{P} \in \Lambda : M\vartheta \vec{P} \neq_\beta I$	$\forall \vartheta : \text{Var} \rightarrow \Lambda, P \in \Lambda : M\vartheta \neq_\beta \lambda x.P$	$M \rightarrow_\beta N \Rightarrow \exists P, Q \in \Lambda : N \rightarrow_\beta (\lambda x.P)Q$
Complement	Solvable: $\exists \vartheta : \text{Var} \rightarrow \Lambda, \vec{P} \in \Lambda : M\vartheta \vec{P} =_\beta I$	Functional: $\exists \vartheta : \text{Var} \rightarrow \Lambda, P \in \Lambda : M\vartheta =_\beta \lambda x.P$	Root-stable: $\exists Z \in \mathfrak{o}_0 - \mathfrak{o}_1, P \in \Lambda : M =_\beta ZP$
Syntactic Characterization	No head normal form	No weak head normal form	No top normal form
Infinitary Semantics	$\text{BT}(M) = \perp$	$\text{LLT}(M) = \perp$	$\text{BeT}(M) = \perp$
Closure Properties	Substitution Left application Abstraction	Substitution Left application	Substitution

 FIGURE 1 Comparing unsolvable terms, terms in \mathfrak{Z} and mute terms.

DEFINITION 5.3

We denote by \mathfrak{Z} the subset of Λ consisting of terms M such that, for all substitutions ϑ and terms N with $M\vartheta \rightarrow_\beta N$, N is not of the form $\lambda x.N'$.

The elements of \mathfrak{Z} are sometimes called ‘zero terms’, but this name has sometimes been applied in the literature to terms having weaker properties. The interested reader is invited to consult the subsection below for a discussion.

5.2 A terminological aside: \mathfrak{Z}

Terms of order 0 are, by definition, terms that cannot be converted to a lambda abstraction. Historically, these terms have sometimes been called *zero terms* [5, 21]. At other times, the expression ‘zero terms’ has been used, even by the same authors, to refer to the class of *unsolvable* terms of order zero. Moreover, this usage is apparently becoming popular, with a number of active researchers employing ‘zero terms’ in this restricted sense [3, 6].

The meaning of the expression ‘zero term’ is therefore disputed, and to avoid ambiguity, we will eschew this term altogether. Yet, we do find that the terminological shift has a decent motivation—especially, with the advent of the infinitary λ -calculus—and shall now briefly comment on it.

Recall that the three canonical infinitary semantics of the λ -calculus are based on BT, Lévy-Longo trees (LLT) and Berarducci trees. These semantics are obtained by coinductively quotienting terms by a chosen subset, the elements of which are deemed to be ‘meaningless’—similarly to quotienting an algebraic structure by some ideal. These sets, respectively, are as follows.

Unsolvable terms: M is *solvable* if, for some substitution $\vartheta : \text{Var} \rightarrow \Lambda$, and some terms \vec{P} , $M\vartheta \vec{P} =_\beta I$. M is *unsolvable* if it is not solvable.

Unsolvable terms of order 0: These are elements of the set \mathfrak{Z} defined in Definition 5.3. It is easy to check that the following are equivalent:

- $M \in \mathfrak{Z}$,
- $M\vartheta \notin \mathfrak{o}_1$ for all $\vartheta : \text{Var} \rightarrow \Lambda$,
- M has no weak hnf.

Mute terms: M is *mute* if every reduct of M reduces to a β -redex. Equivalently, M has no top normal form (i.e. it is a root-active term) [28].

The relationships between these sets are summarized in Figure 1.

Intuitively, one thinks of elements of \mathfrak{Z} as terms that are not convertible to a lambda abstraction (i.e. terms of order 0), which would make the terminology ‘zero terms’ appropriate. The subtlety is that terms of order 0 are not closed under substitution. Indeed, a more robust notion is obtained by defining zero terms to be terms that are *hereditarily* of order 0 (in the sense that all their instances are such). In such an interpretation, zero terms will be precisely the elements of \mathfrak{Z} .

5.3 Fixed points of elements of \mathfrak{Z}

We first prove the proposition below.

PROPOSITION 5.4

If F belongs to \mathfrak{Z} , then $\text{Fix}(F)$ is infinite. Moreover, if F is closed then $\text{Fix}^o(F)$ is infinite as well.

PROOF. Let, for $n \geq 0$, $X_n = Y_n F$, where $Y_n = \Theta_{c_n}$ are pairwise β -distinct fpc’s by Lemma 4.1. Observe that $\{[X_n]_\beta \mid n \in \mathbb{N}\} \subseteq \text{Fix}(F)$. Moreover, when F is closed, then so is $Y_n F$ and hence $[X_n]_\beta$ is an element of $\text{Fix}^o(F)$. The remainder of the proof is devoted to showing the claim below, from which the main result immediately follows.

CLAIM 1 For $m \neq n$, $X_n \neq_\beta X_m$. □

SUBPROOF. The proof uses Claim 2, proved below.

Suppose that $X_n =_\beta X_m$. By the Church–Rosser theorem, there is a λ -term X such that $X_n \rightarrow_\beta X$ and $X \rightarrow_\beta X_m$, and by Claim 2 we obtain

$$Y_n F \rightarrow_\beta F_0(F_1(\cdots Z F_k)) = X = F'_0(F'_1(\cdots Z' F'_{k'})) \rightarrow_\beta Y_m F.$$

We posit that $k = k'$. For contradiction and without loss of generality, assume that $k < k'$. Then we have

$$Z = F'_k, \quad F_k = F'_{k+1}(\cdots Z' F'_{k'}),$$

which contradicts that $F =_\beta F'_k$ belongs to \mathfrak{Z} , while $Y_n =_\beta Z$ is an fpc.

Hence, $k = k'$, but then X has at depth $k + 1$ the subterm $Z = Z'$, which is a β -reduct of both Y_n and Y_m . This is impossible by Lemma 4.1, unless $n = m$. □

CLAIM 2

For any $n \geq 0$ and any λ -term X such that $Y_n F \rightarrow_\beta X$, there is a $k \geq 0$ and there are λ -terms F_0, F_1, \dots, F_k, Z with $F \rightarrow_\beta F_0, F \rightarrow_\beta F_1, \dots, F \rightarrow_\beta F_k$ and $Y_n \rightarrow_\beta Z$ such that $X = F_0(F_1(\cdots Z F_k))$ (intuitively, k is the number of ‘unfoldings’ of the fpc Y_n applied to F).

SUBPROOF. Since Y_n is a reducing fpc, we may consider the infinite reduction sequence

$$Y_n F \rightarrow_\beta F(Y_n F) \rightarrow_\beta F(F(Y_n F)) \rightarrow_\beta \cdots$$

Notice that in any reduction sequence starting from $Y_n F$ there can only be one reduction step contracting a redex that occurs at the root. Indeed, since we are considering $Y_n = \Theta_{c_n} = \mathbf{V}\mathbf{V}c_n$ a redex is created at the root only if it is of shape $(\lambda x.x \Theta_{c_n}^x) F_0$ with $\Theta_{c_n} x \rightarrow_\beta \Theta_{c_n}^x$ and $F \rightarrow_\beta F_0$. Its contractum will therefore have shape $F_0 \Theta_{c_n}^{F_0}$ and none of its descendants will have a redex at the root since $F \in \mathfrak{Z}$ entails that F_0 never reduces to an abstraction. Similarly, in any reduction sequence of this kind there is at most one reduction step contracting a redex occurring at a position of depth k

in the right spine of the syntax tree; this deeper redex can be created only once all redexes at previous positions in the spine have been contracted (those reduction steps correspond to steps in the fixed point combinator unfolding).

Assume wlog that the reduction sequence $Y_n F \rightarrow_\beta X$ contracts $k \geq 0$ redexes in the rightspine of the syntax tree of Y_n . Consider the projection of $Y_n F \rightarrow_\beta X$ across $Y_n F \rightarrow_\beta F(F(\dots F(Y_n F)))$ ($k + 1$ F 's) and write the projection diagram as

$$\begin{array}{ccc} Y_n F & \xrightarrow{\beta} & F(F(\dots F(Y_n F))) \\ \beta \downarrow & & \downarrow \beta \\ X & \xrightarrow{\beta} & H \end{array}$$

By the above arguments, the reduction $X \rightarrow_\beta H$ consist solely of steps inside descendants of F and Y_n ; whence, $X = F_0(F_1(\dots ZF_k))$ for λ -terms Z, F_0, \dots, F_k with $Y_n \rightarrow_\beta Z, F \rightarrow_\beta F_0, F \rightarrow_\beta F_1, \dots, F \rightarrow_\beta F_k$, as desired. \square

The result now follows, as $\{[X_n]_\beta \mid n \in \mathbb{N}\}$ is infinite by Claim 1.

5.4 Recurrent elements of \mathfrak{F} as fixed points

Recall that a λ -term M is *recurrent* if, for all λ -terms N , $M \rightarrow_\beta N$ implies $N \rightarrow_\beta M$. For example, Ω and $\Theta \mathbf{I}$ are recurrent elements of \mathfrak{F} , $\lambda y.y(\Theta \mathbf{I})$ is recurrent, but does not belong to \mathfrak{F} , and $\Omega_3 = \Delta_3 \Delta_3$ is an element of \mathfrak{F} , but is not recurrent.

We proceed to prove a general result that recurrent terms belonging to \mathfrak{F} can only be fixed points of a combinator if *all* λ -terms are fixed points of that combinator, unless the combinator is constant. We first prove Lemma 5.6 below; the general result is Theorem 5.7.

The proofs of both lemma and theorem make use of a result colloquially called ‘Barendregt’s Lemma’; we use it in the following form due to van Daalen (see, e.g. [11] for a comprehensive treatment).

LEMMA 5.5 (Barendregt’s lemma).

Let $M[L/x] \rightarrow_\beta N$. Then there exist a k -hole context $C[x_1, \dots, x_k]$ (with $k \geq 0$), λ -terms $xP_1^1 \dots P_{m_1}^1, \dots, xP_1^k \dots P_{m_k}^k$ with $x \notin \text{FV}(C[\])$ and Q_1, \dots, Q_k such that

- (i): $M \rightarrow_\beta C[xP_1^1 \dots P_{m_1}^1, \dots, xP_1^k \dots P_{m_k}^k]$.
- (ii): $(xP_1^i \dots P_{m_i}^i)[L/x] \rightarrow_\beta Q_i$.
- (iii): $N = C[Q_1, \dots, Q_k]$.

LEMMA 5.6

If $R = C[Q]$ is a recurrent term belonging⁵ to \mathfrak{F} , and $R \rightarrow_\beta Q$, then either $C[z] \rightarrow_\beta R$ or $C[z] \rightarrow_\beta z$ (for $z \notin \text{FV}(C[\])$).

⁵In fact, it is easy to see that this lemma holds for all recurrent R , not just members of \mathfrak{F} . This is because any recurrent term can be presented as $R = N[R_1, \dots, R_k]$, where $N[x_1, \dots, x_k]$ is a normal context (no redexes), and R_i are *root recurrent* (recurrent and reducing to a redex). (This normal form for recurrent terms is obtained by induction on the term structure of R .)

If we now have $R = C[Q] \rightarrow_\beta C[R]$, with $R = N[R_1, \dots, R_k]$ and N normal, then $C[R] = N[R'_1, \dots, R'_k]$ and so $C[N[\bar{R}]] = N[\bar{R}']$. This can only happen if $C[x] = x$, $C[x] = N[\bar{R}']$, or $N[\bar{R}] = R_i$ for some i —in which case $R_i = R \rightarrow_\beta C[R_i]$ and our lemma applies. Since we will not need this level of generality, we do not pursue this observation further.

PROOF. Let $z \notin \text{FV}(C[])$ and assume, for purposes of contradiction, that neither $C[z] \rightarrow_\beta R$, nor $C[z] \rightarrow_\beta z$. Assume now $C[z] \rightarrow_\beta N$. If $z \notin \text{FV}(N)$, we have that $R = C[Q] \rightarrow_\beta N[Q/z] = N$, and by recurrence of R , that $N \rightarrow_\beta R$ and hence $C[z] \rightarrow_\beta R$, contradicting the assumptions. Hence, we must have $z \in \text{FV}(N)$ and, since N is an arbitrary reduct of $C[z]$ and we have assumed that $C[z]$ does not β -reduce to z , every reduct of $C[z]$ must contain z as a free variable strictly below the root.

As R is recurrent and $R \rightarrow_\beta Q$, we have $Q \rightarrow_\beta R$. Thus, we have the reduction sequence

$$R = C[Q] \rightarrow_\beta C[R] = C[C[Q]] \rightarrow_\beta C[C[R]] \rightarrow_\beta C[C[C[R]]] \rightarrow_\beta \dots$$

Hence, for all $n \geq 1$ we have $R \rightarrow_\beta C^n[R]$, and thus by recurrence of R that $C^n[R] \rightarrow_\beta R$. Observe that for every $n \geq 1$, the λ -term $C^n[R]$ is an element of \mathfrak{J} as it is a reduct of R .

CLAIM 3 Let $n \geq 0$ and assume $C^n[R] \rightarrow_\beta W$. Then the length of the longest position in W is at least n . □

SUBPROOF. Proceed by induction:

- $n = 0$: Trivial.
- $n \geq 1$: By Barendregt's lemma, we have that $C[C^{n-1}[R]] \rightarrow_\beta W$ implies the existence of a k -hole context $D[x_1, \dots, x_k]$ (with $k \geq 0$) together with λ -terms $zP_1^1 \dots P_{m_1}^1, \dots, zP_1^k \dots P_{m_k}^k$ and Q_1, \dots, Q_k such that z does not occur free in $D[]$ and the following hold:
 - (1) $C[z] \rightarrow_\beta D[zP_1^1 \dots P_{m_1}^1, \dots, zP_1^k \dots P_{m_k}^k]$,
 - (2) $(zP_1^i \dots P_{m_i}^i)[C^{n-1}[R]/z] \rightarrow_\beta Q_i$ for all $1 \leq i \leq k$, and
 - (3) $W = D[Q_1, \dots, Q_k]$.

As $C^{n-1}[R]$ is an element of \mathfrak{J} , hence cannot reduce to an abstraction, there is a k -hole context $D'[x_1, \dots, x_k]$ such that we may write (i), (ii) and (iii) above as (i) $C[z] \rightarrow_\beta D'[z, \dots, z]$, (ii) $C^{n-1}[R] \rightarrow_\beta Q_i$ for all $1 \leq i \leq k$ and (iii) $W = D'[Q_1, \dots, Q_k]$. Note that by the previous observations, $D'[z, \dots, z]$ cannot have a variable at the root as otherwise $C[z] \rightarrow_\beta z$, a contradiction. Moreover, we cannot have $k = 0$ because, as shown earlier, every reduct of $C[z]$ must contain z as a free variable and z does not occur in $D'[]$. Hence, the length of the longest position in $D'[Q_1, \dots, Q_k]$ is at least one more than the length of the longest position in any of the Q_i 's.

But as $C^{n-1}[R] \rightarrow_\beta Q_i$ for all $1 \leq i \leq k$, the induction hypothesis furnishes that the longest position in any Q_i is at least $n - 1$; hence, the length of the longest position in W is at least n . □

Let $d \geq 1$ be an integer strictly greater than the length of the longest position in R . By Claim 3 above, $C^d[R] \rightarrow_\beta R$ implies that the length of the longest position in R is at least d , a contradiction. Hence, the original assumption leads to contradiction, and we must thus have either $C[z] \rightarrow_\beta R$, or $C[z] \rightarrow_\beta z$, as desired.

THEOREM 5.7

Let F be any λ -term. If there is a recurrent $R \in \mathfrak{J}$ such that $R \in_{\lambda\beta} \text{Fix}(F)$, then the following hold:

- (1) For a fresh variable z , either $Fz \rightarrow_\beta R$ or $Fz \rightarrow_\beta z$.
- (2) Either $F =_\beta \mathbf{KR}$ or $F =_\beta \mathbf{I}$.

- (3) In any λ -theory \mathcal{T} , either $\text{Fix}_{\mathcal{T}}(F) = \{[R]_{\mathcal{T}}\}$ or $\text{Fix}_{\mathcal{T}}(F) = \Lambda_{\mathcal{T}}$. Thus, if $F \in \Lambda^o$ then either $\text{Fix}_{\mathcal{T}}^o(F) = \{[R]_{\mathcal{T}}\}$ or $\text{Fix}_{\mathcal{T}}^o(F) = \Lambda_{\mathcal{T}}^o$.

PROOF. First, we observe that (1) implies both (2) and (3).

- (2): If $Fz \rightarrow_{\beta} R$ for a fresh z , then z must be erased in the reduction sequence, which has therefore length at least 1. By the standardization theorem [2, 11.4.7], $Fz \rightarrow_h (\lambda z.C[z])z \rightarrow_{\beta} C[z] \rightarrow_{\beta} R$; hence, $F \rightarrow_h (\lambda z.C[z]) \rightarrow_{\beta} (\lambda z.R) =_{\beta} \mathbf{K}R$.
 If $Fz \rightarrow_{\beta} z$, then F must β -reduce to an abstraction whence the reduction sequence is non-empty. By the standardization theorem, $Fz \rightarrow_h (\lambda z.C[z])z \rightarrow_{\beta} C[z] \rightarrow_{\beta} z$; hence, $F \rightarrow_h (\lambda z.C[z]) \rightarrow_{\beta} (\lambda z.z)$ and $F =_{\beta} \mathbf{I}$.
- (3): Immediate by (2).

The remainder of the proof is devoted to proving (1). By the above observations, this suffices to prove the theorem.

Suppose $FR =_{\beta} R$ for R a recurrent term in \mathfrak{J} . By the Church–Rosser theorem, there is a λ -term N such that $FR \rightarrow_{\beta} N \leftarrow_{\beta} R$. By recurrence of R , we obtain $N \rightarrow_{\beta} R$ and consequently $FR \rightarrow_{\beta} R$. Let $x \notin \text{FV}(F)$ and set $M = Fx$. By Barendregt’s lemma there is a context $C[x_1, \dots, x_k]$ with $x \notin \text{FV}(C[\])$, λ -terms $P_1^1, \dots, P_{m_1}^1, \dots, P_1^k, \dots, P_{m_k}^k$ and Q_1, \dots, Q_k such that

- (i) $M \rightarrow_{\beta} C[xP_1^1 \dots P_{m_1}^1, \dots, xP_1^k \dots P_{m_k}^k]$,
- (ii) $(xP_1^i \dots P_{m_i}^i)[R/x] \rightarrow_{\beta} Q_i$ for all $1 \leq i \leq k$,
- (iii) $R = C[Q_1, \dots, Q_k]$.

Since $R \in \mathfrak{J}$, point (ii) yields that

$$\begin{aligned} (xP_1^i \dots P_{m_i}^i)[R/x] &= R(P_1^i[R/x]) \dots (P_{m_i}^i[R/x]) \\ &\rightarrow_{\beta} Q_i = R_i S_1^i \dots S_{m_i}^i, \end{aligned}$$

where $R \rightarrow_{\beta} R_i$ and $P_j^i[R/x] \rightarrow_{\beta} S_j^i$.

For all i with $1 \leq i \leq k$, consider the one-hole context

$$C_i[z] = C[Q_1, \dots, Q_{i-1}, z\vec{S}^i, Q_{i+1}, \dots, Q_k].$$

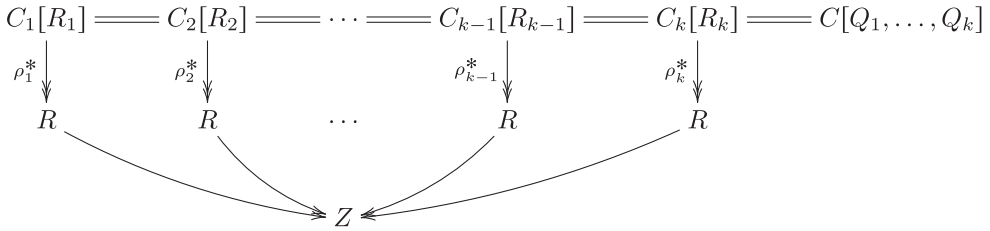
Point (iii) yields that $R = C_i[R_i]$ for each i , so we may apply Lemma 5.6 to conclude that either $C_i[z] \rightarrow_{\beta} z$ or $C_i[z] \rightarrow_{\beta} R$, for each i .

If, for some i , $C_i[z]$ indeed reduces to z , then we conclude the proof by the following sequence of inferences:

1. The vector \vec{S}^i must be empty, so that $m_i = 0$;
2. By the genericity lemma [2, Prop. 14.3.24], $C[Q_1, \dots, Q_{i-1}, z, Q_{i+1}, \dots, Q_k] = C_i[z] \rightarrow_{\beta} z$ implies that $C[x_1, \dots, x_k] \rightarrow_{\beta} x_i$, since $Q_j = R_j \vec{S}^j$ are unsolvable, and z is normal;
3. Hence, $Fx \rightarrow_{\beta} C[xP_1^1, \dots, xP_k^k] \rightarrow_{\beta} x$.

Suppose, on the other hand, that for each i , there is a reduction $\rho_i : C_i[z] \rightarrow_{\beta} R$. We then conclude by the following sequence of inferences.

- (1) For each i , we have the reduction $\rho_i^* = \rho_i[R_i/z] : R = C_i[R_i] \rightarrow_{\beta} R$, which erases the displayed occurrence of R_i along the way.
- (2) By the Church–Rosser theorem, these can be joined together to yield



where all the alternative paths from $C[Q_1, \dots, Q_k]$ to Z are equivalent; hence, no subterm of Z descends from Q_i (which gets erased by ρ_i^*).

- (3) The equivalent composite reductions above therefore lift to a reduction $C[x_1, \dots, x_k] \rightarrow_\beta Z$.
- (4) By recurrence, also $Z \rightarrow_\beta R$.
- (5) By point (i), we get $Fx \rightarrow_\beta C[x_1 \vec{P}^1, \dots, x_k \vec{P}^k] \rightarrow_\beta Z \rightarrow_\beta R$, as desired.

This completes the proof of (i), and of the theorem. \square

The assumptions that the λ -term R is recurrent and belongs to \mathfrak{J} cannot be omitted, as seen in the next example.

EXAMPLE 5.8

Consider the following examples.

1. Let $F = \lambda xy.y(x\mathbf{I})$. Then, $\text{Fix}(F) = \{[\lambda y.yM]_\beta \mid M \in \Lambda, y \notin \text{FV}(M)\}$ and $\text{Fix}^o(F) = \{[\lambda y.yM]_\beta \mid M \in \Lambda^o\}$. Clearly, both $\text{Fix}(F)$ and $\text{Fix}^o(F)$ are infinite and have empty intersection with \mathfrak{J} . Furthermore, both $\text{Fix}(F)$ and $\text{Fix}^o(F)$ contain infinitely many distinct elements $[Q]_\beta$ where Q is a closed recurrent term, namely all λ -terms Q of the form $Q = \lambda y.yR$ where R is a closed recurrent term.
2. Thus, F is a closed λ -term with infinitely many non- β -convertible closed recurrent terms as fixed points, showing that the assumption of $R \in \mathfrak{J}$ in Theorem 5.7 cannot be omitted.
3. Define $\mathbf{J} = \lambda wxy.x(wwxy)$ and note that $\mathbf{JJ}iz \rightarrow_\beta i(\mathbf{JJ}iz)$. Set $F = \mathbf{JJI}$ and, for each $n \geq 0$, consider $Y_n = \Theta_{\mathbf{e}_n}$.
4. Then, as Y_n is an fpc, we get $\{[Y_n F]_\beta \mid n \in \mathbb{N}\} \subseteq \text{Fix}(F)$. Furthermore, by Lemma 4.1 and the construction of F it is easy to see that for $m \neq n$ we have $Y_m F \neq_\beta Y_n F$, whence $\text{Fix}^o(F)$ is infinite. Furthermore, note that $Y_n F \rightarrow_\beta F(Y_n F) \rightarrow_\beta \mathbf{I}(F(Y_n F))$ and that $F(Y_n F)$ does not reduce to $Y_n F$ whence none of the $Y_n F$ is recurrent. It is straightforward to check that for any $n \geq 0$, we have $Y_n F \in \mathfrak{J}$.
5. Hence, F is a closed λ -term with infinitely many non- β -convertible elements of \mathfrak{J} as fixed points, showing that the assumption of R being recurrent in Theorem 5.7 cannot be omitted.

As an application of the previous theorem, recall the notion of *Plotkin terms* from [25]; these are λ -terms P such that, for fresh x , every reduct of Px contains x , and yet $PX =_\beta P\mathbf{I}$ for every closed $X \in \Lambda^o$.

The standard construction of such terms (see [2, Def. 17.3.26]) yields a zero term $Z =_\beta P\mathbf{I}$, which moreover satisfies $PZ =_\beta Z$ (since $Z \in \Lambda^o$). If Z was recurrent, then Theorem 5.7 would apply, implying that P is either identity or constant on all (open) terms. Since P is neither, it follows that Z is not recurrent.⁶

⁶One might suspect that this non-recurrence is due to Plotkin terms being universal generators, but this is not so; the term $WW\mathbf{c}_0$, with $Wwn = \mathbf{K}(ww\mathbf{c}_0)[En, ww(\mathbf{S}^+n)]$ is a universal generator, and it is recurrent.

6 The Fixed Point Property Fails in All Sensible Theories

In this section we prove that no sensible λ -theory \mathcal{T} can satisfy the fixed point property. More precisely, we are going to show that the λ -term defined as

$$\mathcal{E} = \lambda xy.x(x(\mathbf{K}y))\mathbf{\Omega}$$

only has two possible fixed points modulo \mathcal{T} . Interestingly, \mathcal{E} is also a counterexample to the *open* fixed point property. This shows that, in contrast to the theory $\lambda\beta$, neither fixed point properties hold in, say, \mathcal{H} , \mathcal{B} or \mathcal{H}^* .

LEMMA 6.1

$\mathbf{\Omega} \in_{\mathcal{H}} \text{Fix}_{\mathcal{H}}(\mathcal{E})$; hence, $\mathbf{\Omega} \in_{\mathcal{T}} \text{Fix}_{\mathcal{T}}(\mathcal{E})$ for every sensible λ -theory \mathcal{T} .

PROOF. We have $\mathcal{E}\mathbf{\Omega} =_{\mathcal{H}} \lambda y.\mathbf{\Omega}(\mathbf{\Omega}(\mathbf{K}y))\mathbf{\Omega} =_{\mathcal{H}} \mathbf{\Omega}$. □

We now show that the only solvable fixed point of \mathcal{E} in every sensible λ -theory \mathcal{T} is the identity.

PROPOSITION 6.2

Let $M \in \Lambda$ and \mathcal{T} be a sensible λ -theory. If $M \in_{\mathcal{T}} \text{Fix}_{\mathcal{T}}(\mathcal{E})$ then $M \neq_{\mathcal{T}} \mathbf{\Omega}$ entails $M =_{\mathcal{T}} \mathbf{I}$.

PROOF. All the equalities in this proof are intended to take place in the λ -theory \mathcal{T} .

Let $M \neq \mathbf{\Omega}$ be a fixed point of \mathcal{E} in \mathcal{T} . Since M is solvable, it has a hnf:

$$M = \lambda x_0 \dots x_k.x'M_1 \dots M_m. \quad (4)$$

CLAIM 4

The head variable x' of the hnf of M must be x_0 .

SUBPROOF. From $M = \mathcal{E}M$ it follows, for fresh variables y and z , that

$$My = \mathcal{E}My = M(M(\mathbf{K}y))\mathbf{\Omega} = (Mz\mathbf{\Omega})[M(\mathbf{K}y)/z]. \quad (5)$$

Now, let $(y_i)_{i \in \mathbb{N}}$ be fresh variables and denote by σ_i the substitution $[M(\mathbf{K}y_i)/y_{i+1}]$. By iterating equation (5) we get

$$My_0 = (My_1\mathbf{\Omega})\sigma_0 = (My_2\mathbf{\Omega}^{\sim 2})\sigma_1\sigma_0 = \dots = (My_n\mathbf{\Omega}^{\sim n})\sigma_{n-1} \dots \sigma_0. \quad (6)$$

In particular, taking $n = k$, we get

$$\begin{aligned} My_0 &= (My_k\mathbf{\Omega}^{\sim k})\vec{\sigma}^k = x'M_1 \dots M_m[y_k/x_0][\vec{\mathbf{\Omega}}/\vec{x}]\vec{\sigma}^k \\ &= (My_{k+1}\mathbf{\Omega}^{\sim k+1})\vec{\sigma}^{k+1} = x'M_1 \dots M_m[y_{k+1}/x_0][\vec{\mathbf{\Omega}}/\vec{x}]\vec{\sigma}^{k+1}\mathbf{\Omega}, \end{aligned}$$

whence x' cannot be a free variable, for no consistent theory can satisfy $x'\vec{P} = x'\vec{Q}$ with unequal number of P 's and Q 's.

Since M is solvable, so is My_0 , and, by (6), so are $My_n\mathbf{\Omega}^{\sim n}$, for all $n \in \mathbb{N}$.

By Lemma 2.6, we get $x' = x_0$. □

We now need to prove that also the indices k, m must be equal to 0.

CLAIM 5

If $k = 0$ then also $m = 0$.

SUBPROOF. Assume, by contradiction, that $k = 0$ while $m > 0$. On the one hand, we have $M = \lambda x_0.x_0.M_1 \cdots M_m$. On the other hand, we have

$$\begin{aligned} \mathcal{E}M &= \lambda y.M(M(\mathbf{K}y))\Omega \\ &= \lambda y.M(\mathbf{K}yM'_1 \cdots M'_m)\Omega && \text{for } M'_i = M_i[\mathbf{K}y/x_0] \\ &= \lambda y.\mathbf{K}yM'_1 \cdots M'_mM''_1 \cdots M''_m\Omega && \text{for } M'_i = M_i[\mathbf{K}yM'_1 \cdots M'_m/x_0] \\ &= \lambda y.yM'_2 \cdots M'_mM'_1 \cdots M''_m\Omega && \text{as we assumed } m > 0. \end{aligned}$$

Since $M = \mathcal{E}M$ we must have $m = 2m$, which is impossible for $m > 0$. \square

CLAIM 6

If $k = m$ then $k = 0$.

SUBPROOF. By induction on $k \in \mathbb{N}$, we show that $M = \lambda x_0 \dots x_k.x_0.M_1 \cdots M_k$ implies $M = \mathbf{I}$.

$k = 0$: Trivial, since M has already the required form.

$k > 0$: In the induction case, we have the following chain of equalities:

$$\begin{aligned} M &= \mathcal{E}M && \text{as } M \in \mathcal{T} \text{ Fix}_{\mathcal{T}}(\mathcal{E}) \\ &= \lambda y.M(M(\mathbf{K}y))\Omega && \text{by def. of } \mathcal{E} \\ &= \lambda y.M(\lambda x_1 \dots x_k.\mathbf{K}yM'_1 \cdots M'_k)\Omega && \text{for } M'_i = M_i[\mathbf{K}y/x_0] \\ &=_{\beta} \lambda y.M(\lambda x_1 \dots x_k.yM'_2 \cdots M'_k)\Omega && \text{since } k > 0 \\ &= \lambda y.(\lambda w_0 \dots w_k.w_0N_1 \cdots N_k)(\lambda x_1 \dots x_k.yM'_2 \cdots M'_k)\Omega && \text{by } \alpha\text{-renaming } M \\ &=_{\beta} \lambda yw_2 \dots w_k.((\lambda x_1 \dots x_k.yM'_2 \cdots M'_k)N'_1 \cdots N'_k)[\Omega/w_1] && \\ &\quad \text{for } N'_j = N_j[\lambda x_1 \dots x_k.yM'_2 \cdots M'_k/w_0] && \\ &=_{\beta} \lambda yw_2 \dots w_k.yM''_2 \cdots M''_k, && \\ &\quad \text{where } M''_i = M'_i[N'_1/x_1] \cdots [N'_k/x_k][\Omega/w_1] && \\ &= \lambda z_0 \dots z_{k-1}.z_0P_1 \cdots P_{k-1} && \text{by } \alpha\text{-renaming} \\ &= \mathbf{I} && \text{by ind. hyp.} \end{aligned}$$

Since $M = \lambda x_0.x_0$, we conclude that $k = 0$. \square

Assume now $k > 0$ and $k \neq m$ towards a contradiction. Easy calculations give

$$My\Omega = \lambda x_2 \dots x_k.y(M_1[y/x_0][\Omega/x_1]) \cdots (M_m[y/x_0][\Omega/x_1]).$$

As a matter of notation we set $V = \lambda y.My\Omega$, and to simplify the reasoning on the indices we perform some α -renaming, namely we let

$$V = \lambda yz_1 \dots z_{k-1}.yV_1 \cdots V_m,$$

where $V_i = M_i[y/x_0][\Omega/x_1][z_1/x_2] \cdots [z_{k-1}/x_k]$ for $1 \leq i \leq m$. We first prove the following claims.

CLAIM 7

For all $n \in \mathbb{N}$, we have $My = V^n(M(\mathbf{K}^ny))$.

SUBPROOF. We proceed by induction on n .

- Case $n = 0$. Trivial since $My = M(\mathbf{K}^0y) = V^0(M(\mathbf{K}^0y))$.

- Case $n + 1$. We have

$$\begin{aligned}
 My &= \mathcal{E} My && \text{as } M \in \mathcal{T} \text{ Fix}_{\mathcal{T}}(\mathcal{E}) \\
 &= M(M(\mathbf{K}_y))\Omega && \text{by def. of } \mathcal{E} \\
 &= V(M(\mathbf{K}_y)) && \text{by def. of } V \\
 &= V(V^n(M(\mathbf{K}^n(\mathbf{K}_y)))) && \text{by induction hypothesis} \\
 &= V^{n+1}(M(\mathbf{K}^{n+1}y)) &&
 \end{aligned}$$

□

In the proofs below we use the following basic properties of \mathbf{K} (for a fresh x):

- (K₁) for all $i, j \geq 0$ we have $\lambda w_1 \dots w_i. \mathbf{K}^j x =_{\beta} \mathbf{K}^{i+j} x$,
- (K₂) if $i > j$ then $(\mathbf{K}^i x) P_1 \dots P_j =_{\beta} \mathbf{K}^{i-j} x$ for arbitrary $P_1, \dots, P_j \in \Lambda$,
- (K₃) if $i \leq j$ then $(\mathbf{K}^i x) P_1 \dots P_j =_{\beta} x P_{i+1} \dots P_j$ for arbitrary $P_1, \dots, P_j \in \Lambda$.

CLAIM 8

For all $n \geq m$, we have $My = V(V^n(\mathbf{K}^{n+1-m+k}y))$.

SUBPROOF. We establish the following chain of equalities:

$$\begin{aligned}
 My &= V^{n+1}(M(\mathbf{K}^{n+1}y)) && \text{by Claim 7} \\
 &= V^{n+1}(\lambda x_1 \dots x_k. (\mathbf{K}^{n+1}y) M'_1 \dots M'_m) && \text{by (4) with } x' = x_0 \\
 &\quad \text{where } M'_i = M_i[\mathbf{K}^{n+1}y/x_0] && \text{for } 1 \leq i \leq m \\
 &=_{\beta} V^{n+1}(\lambda x_1 \dots x_k. \mathbf{K}^{n+1-m}y) && \text{by (K}_2\text{), since } n+1 > m \\
 &=_{\beta} V^{n+1}(\mathbf{K}^{n+1-m+k}y) && \text{by (K}_1\text{)}.
 \end{aligned}$$

□

We split into subcases, depending on whether m is greater than k .

CLAIM 9

When $k > m$ we have for all $n \in \mathbb{N}$ (and for appropriate $X_i \in \Lambda$):

- (i) $V^n(Vy) = \lambda x_1 \dots x_{k-1+(k-1-m)n}. y X_1 \dots X_m$,
- (ii) if $n \geq m$ then $My = \mathbf{K}^{(2+n)(k-m)}y$.

SUBPROOF. (i) We proceed by induction on n .

- If $n = 0$ then the case follows by definition of V .
- If $n > 0$ then we have

$$\begin{aligned}
 V^n(Vy) &= V(V^{n-1}(Vy)) && \text{by def.} \\
 &= V(\lambda x_1 \dots x_{k-1+(k-1-m)(n-1)}. y X_1 \dots X_m) && \text{by ind. hyp.} \\
 &=_{\beta} \lambda z_1 \dots z_{k-1}. (\lambda x_1 \dots x_{k-1+(k-1-m)(n-1)}. y \vec{X}) V'_1 \dots V'_m, \\
 &\quad \text{where } V'_i = V_i[\lambda x_1 \dots x_{k-1+(k-1-m)(n-1)}. y \vec{X}/y] && \text{for } 1 \leq i \leq m \\
 &=_{\beta} \lambda z_1 \dots z_{k-1} x_{m+1} \dots x_{k-1+(k-1-m)(n-1)}. y X'_1 \dots X'_m && \text{as } k > m, \\
 &\quad \text{where } X'_i = X_i[V'_1/x_1] \dots [V'_m/x_m]
 \end{aligned}$$

So the number of abstractions is $k - 1 + k - 1 + (k - 1 - m)(n - 1) - m = k - 1 + (k - 1 - m)n$.

(ii) For $n \geq m$ we have the following:

$$\begin{aligned}
 My &= V^n(V(\mathbf{K}^{n+1-m+k}y)) && \text{by Claim 8} \\
 &= \lambda x_1 \dots x_{k-1+(k-1-m)n}. (\mathbf{K}^{n+1-m+k}y) X_1 \dots X_m && \text{by (i)} \\
 &=_{\beta} \lambda x_1 \dots x_{k-1+(k-1-m)n}. \mathbf{K}^{(n+1-m+k)-m}y && \text{by (K}_2\text{) as } n \geq m, k > m \\
 &=_{\beta} \mathbf{K}^{k-1+(k-1-m)n+(n+1-m+k)-m}y && \text{by (K}_1\text{)}.
 \end{aligned}$$

So, the number of \mathbf{K} 's is $k - 1 + (k - 1 - m)n + (n + 1 - m + k) - m = (k - 1 - m)n + n + 2k - 2m = (k - 1 - m + 1)n + 2(k - m) = (k - m)(n + 2)$. □

854 *The fixed point property and a technique to harness double fixed point combinators*

In Claim 9(ii) we have shown that, for all n large enough, My has a hnf with $(k-m)(n+2)$ external λ -abstractions and 0 applications. By Remark 2.9, we have $(k-m)(n+2) = k-m$ for all such n , which is only possible if this quantity is independent from n . As we are supposing $k > m$ this is impossible.

CLAIM 10

When $0 < k < m$ we have for all $n \in \mathbb{N}$ (and for appropriate $X_i \in \Lambda$):

- (i) $V^n(Vy) = \lambda x_1 \dots x_{k-1}.yX_1 \dots X_{m+(m-k+1)n}$,
- (ii) if $n \geq m$ then $My = \lambda x_1 \dots x_{k-1}.yX_1 \dots X_{(m-k)n+2m-k-1}$.

SUBPROOF. (i) We proceed by induction on n .

- If $n = 0$ then the case follows by definition of V .
- If $n > 0$ then we have

$$\begin{aligned}
 V^n(Vy) &= V(V^{n-1}(Vy)) && \text{by def.} \\
 &= V(\lambda x_1 \dots x_{k-1}.yX_1 \dots X_{m+(m-k+1)(n-1)}) && \text{by ind. hyp.} \\
 &\equiv_{\beta} \lambda \vec{z}.(\lambda x_1 \dots x_{k-1}.yX_1 \dots X_{m+(m-k+1)(n-1)})V_1 \dots V_m \\
 &\equiv_{\beta} \lambda z_1 \dots z_{k-1}.yX'_1 \dots X'_{m+(m-k+1)(n-1)}V_k \dots V_m && \text{as } k < m, \\
 &\quad \text{where } X'_i = X_i[V_1/x_1] \dots [V_{k-1}/x_{k-1}] && \text{for } 1 \leq i \leq m.
 \end{aligned}$$

So the number of applications is $m + (m-k+1)(n-1) + m-k+1 = m + (m-k+1)n$.

(ii) For $n \geq m$ we have the following:

$$\begin{aligned}
 My &= V^n(V(\mathbf{K}^{n+1-m+k}y)) && \text{by Claim 8} \\
 &= \lambda x_1 \dots x_{k-1}.(\mathbf{K}^{n+1-m+k}y)X_1 \dots X_{m+(m-k+1)n} && \text{by (i)} \\
 &\equiv_{\beta} \lambda x_1 \dots x_{k-1}.yX_{(n+1-m+k)+1} \dots X_{m+(m-k+1)n},
 \end{aligned}$$

where the last equality follows by (K_3) since $k < m \leq n$ so that $m + (m-k+1)n - (n+1-m+k) = (m-k+1)n + m - n - 1 + m - k = (m-k+1-1)n + 2m - k - 1 = (m-k)n + 2m - k - 1 > 0$. In particular, the number of applications is what is claimed. \square

By Claim 10(ii), for all n large enough, My has a hnf with $(m-k)n + 2m - k - 1$ applications and $k-1$ external abstractions, so the difference is $(m-k)n + 2m - 2k$. By Remark 2.9, we must have $(m-k)n + 2m - 2k = m - k$ for all such n , which is only possible if this quantity is independent from n . As we are supposing $k < m$ this is impossible.

As we ruled out all other possibilities, we conclude $k = m = 0$ and $M = \mathbf{I}$.

As a consequence of Lemma 6.1 and Proposition 6.2 we obtain the following.

COROLLARY 6.3

For every sensible λ -theory \mathcal{T} , $\text{Fix}_{\mathcal{T}}^o(\mathcal{E}) = \text{Fix}_{\mathcal{T}}(\mathcal{E}) = \{[\mathbf{Q}]_{\mathcal{T}}, [\mathbf{I}]_{\mathcal{T}}\}$ is of cardinality 2.

We are now able to present the main result of the paper.

THEOREM 6.4

No sensible λ -theory \mathcal{T} satisfies the fixed point property.

This gives a partial answer to Problem 3 and has the following corollary.

COROLLARY 6.5

The λ -theory \mathcal{B} satisfies the range property, but not the fixed point property.

We conclude this section with one more observation.

PROPOSITION 6.6

Let \mathcal{T} be a sensible λ -theory. For all $k > 0$, there exists $M_k \in \mathcal{A}^o$ such that $\text{Fix}_{\mathcal{T}}(M_k) = \text{Fix}_{\mathcal{T}}^o(M_k)$ has cardinality k .

PROOF. We define inductively the following sequence of terms:

$$F_1 = \mathbf{F} = \lambda xy.y, \quad F_2 = \mathbf{E}, \quad F_{n+1} = \lambda x.[\mathbf{E}(\pi_1 x), \pi_1 x F_n(\pi_2 x)] \text{ for } n \geq 2$$

and proceed by induction on k .

The case $k = 1$ is trivial since $\text{Fix}_{\mathcal{T}}(\mathbf{F}) = \text{Fix}_{\mathcal{T}}^o(\mathbf{F}) = \{[\mathbf{I}]\mathcal{T}\}$.

The case $k = 2$ follows by Proposition 6.2.

Assume $k > 2$. Suppose that $X \in \text{Fix}_{\mathcal{T}}(F_k)$, which means that $X =_{\mathcal{T}} F_k X$. Then X must be such that $X =_{\mathcal{T}} [X_1, X_2]$, where

$$X_1 =_{\mathcal{T}} \mathbf{E} X_1 \quad X_2 =_{\mathcal{T}} X_1 F_{k-1} X_2.$$

Since $X_1 =_{\mathcal{T}} \mathbf{E} X_1$, Proposition 6.2 entails that either $X_1 =_{\mathcal{T}} \mathbf{I}$ or $X_1 =_{\mathcal{T}} \mathbf{I}$. In the former case we must have also $X_2 =_{\mathcal{T}} \mathbf{I}$. In the latter, the fact that $X_1 =_{\mathcal{T}} \mathbf{I}$ entails that $X_2 =_{\mathcal{T}} F_{k-1} X_2$. By induction hypothesis, there are exactly $k - 1$ solutions to this equation (modulo \mathcal{T}). It is easy to check that each of these solutions indeed furnishes a fixed point of F_k . Therefore, the set

$$\text{Fix}_{\mathcal{T}}(F_k) = \{[\mathbf{I}, \mathbf{I}]\mathcal{T}\} \cup \{[\mathbf{I}, X]\mathcal{T} \mid X \in_{\mathcal{T}} \text{Fix}_{\mathcal{T}}(F_{k-1})\}$$

consists of closed terms and, by Remark 2.9, has cardinality k . \square

7 The Double Fixed Points Problem

In this section we focus on Problem 1, originally stated by Statman [29] and attacked by Intrigila [15], namely the question of whether double fixed point combinators exist. Intrigila's proposal is centred on the remark that, in the BT model, both Y and $Y\delta$ are indeed equated and thus that somehow fixed point unrollings had to be tamed with. While Intrigila defined a notion of weight to perform this task, we approach the question differently by factoring the behaviour of the fixed point combinator itself through a notion of interpretation of the λY -calculus in the λ -calculus and the identification of structural properties of this interpretation from which the non-existence of double fixed point combinators would follow.

7.1 Background on the λY -calculus

The λY -calculus is an extension of the untyped λ -calculus with a unary term constructor Y representing a fixed point combinator. Formally, the set \mathcal{A}_Y of λY -terms is generated by the following grammar:

$$\mathcal{A}_Y : \quad M, N ::= x \mid MN \mid \lambda x.M \mid YM.$$

In order to endow the Y construct with the behaviour of a fixed point combinator, we consider an additional reduction \rightarrow_Y , which is the contextual closure of the rule:

$$YM \rightarrow M(YM). \quad Y$$

The λY -calculus thus becomes a higher-order rewriting system with reduction $\rightarrow_{\beta Y}$ generated by the rules (β) and (Y) . Most of the notions introduced in Section 2 for the λ -calculus are inherited by

the λY -calculus in the obvious way. In particular, a λY -theory is a congruence on Λ_Y containing the βY -conversion.

Several standard references provide background on the λY -calculus [1, 24, 30]. The usual rewriting theoretic properties of the λ -calculus carry over to the λY -extension with virtually the same proofs. We still review these arguments as later on we will employ some refinements of them, but we refer to Appendix A for the most technical proofs.

THEOREM 7.1

The reduction $\rightarrow_{\beta Y}$ is confluent.

PROOF. The λY -calculus possesses two rewriting rules. By inspection, it is evident that the system is orthogonal—there is no possible overlap between redex patterns of the two rules. We conclude since, by [4, Thm. 11.6.19], every orthogonal higher-order term rewriting system is confluent. \square

As a consequence, two βY -convertible λY -terms M and N have a common reduct.

COROLLARY 7.2

Let $M, N \in \Lambda_Y$. If $M =_{\beta Y} N$, then there exists $Z \in \Lambda_Y$ such that $M \rightarrow_{\beta Y} Z \beta Y \leftarrow N$.

In fact, the system λY is a conservative extension of the λ -calculus.

COROLLARY 7.3

λY is conservative over λ .

PROOF. Let $M, N \in \Lambda$ such that $M =_{\beta Y} N$. By Corollary 7.2, there is a λY -term Z such that $M \rightarrow_{\beta Y} Z \beta Y \leftarrow N$. Since neither M nor N contain the symbol Y , and this symbol cannot be created by β -reduction, there is no point during these reductions where such a symbol can appear. Consequently, there is no point during these reductions where the Y -rule can be applied. We conclude that these reductions in λY are actually reductions in λ , hence $M =_{\beta} N$ holds. \square

7.1.1 Standardization and parallel reduction We now present some reduction relations that are well known in the setting of the λ -calculus, and are here extended to the λY -calculus.

DEFINITION 7.4

- (1) The *weak head reduction* is defined by the following two rules (for $k \geq 0$):

$$\begin{aligned} (\lambda x.M)N_0 \cdots N_k &\rightarrow_w M[N_0/x]N_1 \cdots N_k \\ YN_0 \cdots N_k &\rightarrow_w N_0(YN_0)N_1 \cdots N_k \end{aligned}$$

- (2) The *standard reduction* is obtained from the weak head reduction by setting:

$$\begin{array}{c} \frac{M \rightarrow_w M' \quad M' \rightarrow_s N}{M \rightarrow_s N} \quad \frac{}{x \rightarrow_s x} \\[10pt] \frac{M \rightarrow_s M'}{YM \rightarrow_s YM'} \quad \frac{M \rightarrow_s M'}{\lambda x.M \rightarrow_s \lambda x.M'} \quad \frac{M \rightarrow_s M' \quad N \rightarrow_s N'}{MN \rightarrow_s M'N'} \end{array}$$

(3) The *parallel reduction* is the least congruence closed under simultaneous development:

$$\frac{M \Rightarrow_p M' \quad N \Rightarrow_p N'}{(\lambda x.M)N \Rightarrow_p M'[N'/x]} \quad \frac{M \Rightarrow_p M' \quad YM \Rightarrow_p N}{YM \Rightarrow_p M'N} \quad \frac{}{x \Rightarrow_p x}$$

$$\frac{M \Rightarrow_p M'}{YM \Rightarrow_p YM'} \quad \frac{M \Rightarrow_p M'}{\lambda x.M \Rightarrow_p \lambda x.M'} \quad \frac{M \Rightarrow_p M' \quad N \Rightarrow_p N'}{MN \Rightarrow_p M'N'}$$

We refer to the appendix for the basic results on these notions of reduction, including the standardization theorem. The proofs in the next two sections will only use the following facts about parallel reduction—whose proofs may be found there as well. Note that the transitive closure of parallel reduction is equal to $\rightarrow_{\beta\Upsilon}$.

PROPOSITION 7.5

For $M, N \in \Lambda_Y$, we have that $M \Rightarrow M'$ and $M' \rightarrow_Y N$ entail $M \Rightarrow N$. In particular, $M \rightarrow_Y N$ implies $M \Rightarrow N$.

PROPOSITION 7.6

For $M, N \in \Lambda_Y$, we have that $M \Rightarrow M'$ and $N \Rightarrow N'$ entail $M[N/x] \Rightarrow M'[N'/x]$.

7.1.2 The simply typed case We now consider the version of λY endowed with simple types over one ground type o . The typing restriction will prove to have several important advantages.

DEFINITION 7.7

The *typed λY -calculus*, λY_{\rightarrow} , is an extension of the simply typed λ -calculus obtained by adding a new unary term constructor Y_A , for each type A :

$$A, B \in \mathbb{T} \quad ::= \quad o \mid A \rightarrow B$$

$$M, N \in \Lambda_Y^{\rightarrow} \quad ::= \quad x \mid MN \mid \lambda x:A.M \mid Y_A M.$$

The typing rule for the new term constructor is the following:

$$\frac{\Gamma \vdash M : A \rightarrow A}{\Gamma \vdash Y_A M : A}$$

The reduction rule is as in the untyped case:

$$Y_A M \rightarrow M(Y_A M). \quad \Upsilon$$

PROPOSITION 7.8

λY_{\rightarrow} satisfies the subject reduction property.

PROOF. Routine. □

7.2 Interpretation of the constructor Y by fixed point combinators

7.2.1 Interpretation by fixed point combinators We have seen that in a λY -term M the constant Y represents a generic fixed point combinator. Therefore, it is possible to retrieve a regular λ -term by substituting some fpc Y for every occurrence of Y in M . The λ -term M' so defined is called the ‘interpretation of M in Λ ’—and it depends on Y . In the next definition we are more liberal and consider also the case where Y is substituted by a weak fixed point combinator.

DEFINITION 7.9

Given a weak fpc $Y \in \Lambda$, we define the *interpretation of a λY -term in Λ with respect to Y* as the map $\llbracket \cdot \rrbracket_Y : \Lambda_Y \rightarrow \Lambda$ given by

$$\begin{aligned}\llbracket x \rrbracket_Y &= x \\ \llbracket MN \rrbracket_Y &= \llbracket M \rrbracket_Y \llbracket N \rrbracket_Y \\ \llbracket \lambda x. M \rrbracket_Y &= \lambda x. \llbracket M \rrbracket_Y \\ \llbracket \Upsilon M \rrbracket_Y &= Y \llbracket M \rrbracket_Y.\end{aligned}$$

Such an interpretation is clearly compositional and enjoys several interesting properties.

LEMMA 7.10 (Substitution lemma for λY).

Let $M, N \in \Lambda_Y$ and let $Y \in \Lambda$ be a weak fpc. Then, for all $x \notin \text{FV}(Y)$, we have

$$\llbracket M[N/x] \rrbracket_Y = \llbracket M \rrbracket_Y \llbracket \llbracket N \rrbracket_Y / x \rrbracket.$$

PROOF. Straightforward by compositionality of the interpretation map $\llbracket \cdot \rrbracket_Y$. \square

In general a weak fpc Y can be such that $Yx \rightarrow_\beta x(Y'x)$ for $Y \neq_\beta Y'$, and in this case the interpretation is unsound; we have $\Upsilon x \rightarrow_Y x(\Upsilon x)$ but $\llbracket \Upsilon x \rrbracket_Y \neq_\beta \llbracket x(\Upsilon x) \rrbracket_Y$. However, when Y is an actual fpc the resulting interpretation is sound.

PROPOSITION 7.11 (Soundness).

Let $Y \in \Lambda$ be an fpc. For all $M, N \in \Lambda_Y$, if $M =_{\beta_Y} N$ then $\llbracket M \rrbracket_Y =_\beta \llbracket N \rrbracket_Y$.

PROOF. First, notice that by Lemma 7.10, if $M =_\beta M'$, then we have $\llbracket M \rrbracket_Y =_\beta \llbracket M' \rrbracket_Y$. Notice also that, if $M =_Y M'$, then we have $\llbracket M \rrbracket_Y =_\beta \llbracket M' \rrbracket_Y$ because Y is an fpc. The result then easily follows by induction on the number of alternations between $=_\beta$ and $=_Y$ in a proof that $M =_{\beta_Y} N$. \square

REMARK 7.12

The converse to the above proposition fails for two reasons. One of these is rather trivial, the other much deeper.

- The first problem has to do with the fact that the interpretation function $\llbracket \cdot \rrbracket_Y$ is not injective even with respect to α -conversion. For example, fix any untyped fpc Y , and consider $M = \lambda x. [Yx, \Upsilon x]$ and $N = \lambda x. [Yx, Yx]$. Trivially $\llbracket M \rrbracket_Y = \llbracket N \rrbracket_Y$, but $M \neq_{\beta_Y} N$ by a Church–Rosser argument.
- The exotic reason is related to the Plotkin terms already discussed on Page 55; there exist (unsolvable) λ -terms $P \in \Lambda^o$ with the property that $PX =_\beta P\mathbf{I}$ for all $X \in \Lambda^o$, and yet $Px \rightarrow_\beta P'$ implies that $x \in \text{FV}(P')$.

For the counterexample now take $M = P\mathbf{I}$ and $N = P(\lambda z. \Upsilon z)$. Just as x can never be erased from Px by any β -reduction, also Υ can never be erased from $P(\lambda z. \Upsilon z)$ by any β_Y -reduction. Yet, for a closed fpc Y , $\llbracket \lambda z. \Upsilon z \rrbracket_Y$ becomes a closed λ -term, and so $\llbracket N \rrbracket_Y =_\beta P\mathbf{I} =_\beta \llbracket M \rrbracket_Y$.

7.2.2 Interpretation of Υ by fpc's in the typed case We now prove that both of the pathologies described in Remark 7.12 disappear when considering the simply typed λY -calculus. We start by showing that the interpretation becomes injective.

DEFINITION 7.13

For a given fpc Y , the interpretation of $\lambda Y \rightarrow$ in Λ is defined as in the untyped case, namely forgetting the types.

PROPOSITION 7.14

Let Y be an fpc. Then the map

$$\llbracket \cdot \rrbracket_Y : \Lambda_Y^{\rightarrow} \rightarrow \Lambda$$

is injective—with respect to syntactic equality.

PROOF. The structure of $\llbracket M \rrbracket_Y$ is completely determined by M ; the only two clauses in the definition of $\llbracket \cdot \rrbracket_Y$, which result in the same term constructor, are those for the application and for Y .

$$\begin{aligned} \llbracket M_1 M_2 \rrbracket_Y &= \llbracket M_1 \rrbracket_Y \llbracket M_2 \rrbracket_Y \\ \llbracket Y N \rrbracket_Y &= Y \llbracket N \rrbracket_Y \end{aligned}$$

Suppose there are $M_1, M_2, N \in \Lambda_Y^{\rightarrow}$ such that

$$\llbracket M_1 \rrbracket_Y \llbracket M_2 \rrbracket_Y = Y \llbracket N \rrbracket_Y$$

Then we must have $\llbracket M_1 \rrbracket_Y = Y$. We claim that this is impossible. First of all, note that Y itself is not a $\lambda Y \rightarrow$ -term, so $M_1 \neq Y$. Now, if Y occurs in M_1 then Y occurs as a strict subterm of $\llbracket M_1 \rrbracket_Y = Y$. This is impossible for finite terms.

Otherwise Y does not occur in M_1 , $\llbracket M_1 \rrbracket_Y = M_1$ and M_1 is a Y -free simply typed term, thus normalizing, which Y is not. \square

PROPOSITION 7.15

Let $M, N \in \Lambda_Y^{\rightarrow}$. Suppose that, for every fpc Y , $\llbracket M \rrbracket_Y =_{\beta} \llbracket N \rrbracket_Y$. Then $M =_{\beta_Y} N$.

PROOF. We consider the interpretation of Y by Turing's fpc Θ and to lighten the notation we simply write $\llbracket \cdot \rrbracket$ for $\llbracket \cdot \rrbracket_{\Theta}$. We first show that for any one-step reduct $\llbracket M \rrbracket \rightarrow_{\beta} Z'$ there exists a $\lambda Y \rightarrow$ -term Z such that

$$M =_{\beta_Y} Z, \quad \llbracket Z \rrbracket = Z'.$$

To see this, write $\llbracket M \rrbracket = C[R]$, where R is the contracted redex. Notice that R cannot be a proper subterm of Θ , which has only one redex, occurring at the root:

$$\Theta = \mathbf{W}\mathbf{W}, \quad \mathbf{W} = \lambda w x. x(wwx).$$

Case 1: If R is indeed the λ -term $\Theta = \mathbf{W}\mathbf{W}$, then it must descend from an occurrence of Y ; in this case we have

$$\begin{aligned} \llbracket M \rrbracket &= C[\Theta] = C'[\Theta \llbracket N \rrbracket], & C[x] &= C'[x \llbracket N \rrbracket] \\ M &= C_0[Y N], & \llbracket C_0[X] \rrbracket &= C'[\llbracket X \rrbracket] \\ \llbracket M \rrbracket \rightarrow_{\beta} Z' &= C'[(\lambda x. x(\mathbf{W}\mathbf{W}x)) \llbracket N \rrbracket]. \end{aligned}$$

But now we have

$$\begin{aligned} M &= C_0[\Upsilon N] \rightarrow_Y C_0[N(\Upsilon N)] \xrightarrow{\beta} C_0[(\lambda x.x(\Upsilon x))N] = Z \\ \llbracket Z \rrbracket &= \llbracket C_0[(\lambda x.x(\Upsilon x))N] \rrbracket \\ &= C'[\llbracket (\lambda x.x(\mathbf{W}\mathbf{W}_x))N \rrbracket] \\ &= C'[(\lambda x.x(\mathbf{W}\mathbf{W}_x))\llbracket N \rrbracket] = Z', \end{aligned}$$

where we find $M =_{\beta_Y} Z$ and $\llbracket Z \rrbracket = Z'$.

Case 2: If R does not come from Υ , then the only possibility left is that it is the image of a redex, which already appears in M :

$$\begin{aligned} \llbracket C_0[X] \rrbracket &= C[\llbracket X \rrbracket], & \llbracket (\lambda x.P)Q \rrbracket &= R \\ M &= C_0[(\lambda x.P)Q] \rightarrow_{\beta} C_0[P[Q/x]] = Z \\ \llbracket M \rrbracket &= \llbracket C_0[(\lambda x.P)Q] \rrbracket = C[R] \rightarrow_{\beta} Z' = C[\llbracket P \rrbracket[\llbracket Q \rrbracket/x]] \\ \llbracket Z \rrbracket &= \llbracket C_0[P[Q/x]] \rrbracket = C[\llbracket P[Q/x] \rrbracket]. \end{aligned}$$

By Lemma 7.10, we have $\llbracket P \rrbracket[\llbracket Q \rrbracket/x] = \llbracket P[Q/x] \rrbracket$; hence, $Z' = \llbracket Z \rrbracket$ and $M \rightarrow_{\beta} Z$.

The result then follows by induction, applying Proposition 7.14. \square

7.3 The reduction extension properties

We now present structural properties of the interpretation map $\llbracket - \rrbracket_Y$ that we call ‘reduction extension properties’. To present them in diagrammatic form, we first need to introduce some notations.

NOTATION 7.16

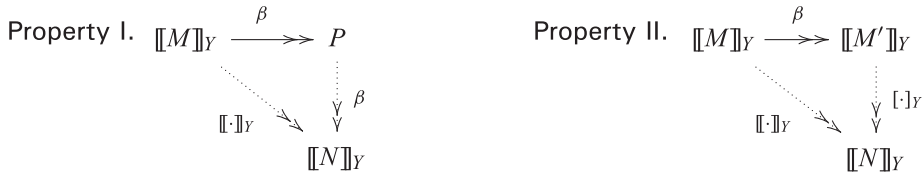
Let $M, N \in \Lambda_Y$, and let Y be a weak fpc. We write

$$\llbracket M \rrbracket_Y \xrightarrow{\llbracket \cdot \rrbracket_Y} \llbracket N \rrbracket_Y$$

whenever $M \rightarrow_{\beta_Y} N$.

DEFINITION 7.17

A weak fpc Y satisfies the *reduction extension properties* if the following properties hold for all $M \in \Lambda_Y$. (Note that $M', N \in \Lambda_Y$ while $P \in \Lambda$.)



where solid arrows denote the assumption reductions and dotted arrows denote the entailed ones. In words, property I states that for all $M \in \Lambda_Y, P \in \Lambda$, $\llbracket M \rrbracket_Y \rightarrow_{\beta} P$ entails that there exists an $N \in \Lambda_Y$ such that $M \rightarrow_{\beta_Y} N$ and $P \rightarrow_{\beta} \llbracket N \rrbracket_Y$. Similarly, property II states that for all $M, M' \in \Lambda_Y$, $\llbracket M \rrbracket_Y \rightarrow_{\beta} \llbracket M' \rrbracket_Y$ entails that there exists an $N \in \Lambda_Y$ such that $M \rightarrow_{\beta_Y} N \xrightarrow{\beta_Y} M'$.

Those properties are interesting because of the following observation.

PROPOSITION 7.18

Let $Y \in \Lambda$ be a weak fpc. If Y satisfies the reduction extension properties then $Y \neq_\beta Y\delta$.

PROOF. Suppose, by way of contradiction, that $Y =_\beta Y\delta$ holds. This entails that $Yx =_\beta Y\delta x$, so by confluence these λ -terms have a common reduct X satisfying $Yx \rightarrow_\beta X \leftarrow_\beta Y\delta x$. Furthermore, by definition of $\llbracket \cdot \rrbracket_Y$, we have $Yx = \llbracket Yx \rrbracket_Y$ and $Y\delta x = \llbracket Y\delta x \rrbracket_Y$. By property I, $Yx \rightarrow_{\beta_Y} M$ for some $M \in \Lambda_Y$ such that $X \rightarrow_\beta \llbracket M \rrbracket_Y$.

Now $\llbracket Y\delta x \rrbracket_Y \rightarrow_\beta X \rightarrow_\beta \llbracket M \rrbracket_Y$ and this entails, by property II, that $Y\delta x \rightarrow_{\beta_Y} N_{\beta_Y} \leftarrow M$. Therefore, $Y\delta x \rightarrow_{\beta_Y} N_{\beta_Y} \leftarrow M_{\beta_Y} \leftarrow Yx$. This is a contradiction, since no β_Y -reduct of Yx contains a δ , while every reduct of $Y\delta x$ contains one—it occurs at the innermost position, with a unique descendant of Y applied to it. \square

REMARK 7.19

- (1) We will prove reduction extension property I for a class of reducing fpc's. We conjecture that this property actually holds for all reducing fpc's, and that our technique will be useful to treat the general case as well.
- (2) Property I can be satisfied by a weak fpc Y satisfying $Yx \rightarrow_\beta x(Y'x)$ even when $Y \neq_\beta Y'$, because we might have $Y'x \rightarrow_\beta x(Yx)$, and $Yx \rightarrow_\beta x^2(Yx)$.
- (3) An fpc Y satisfying property I, cannot satisfy property II for all $M \in \Lambda_Y$, because together they would imply the completeness of the interpretation $[\cdot]_Y$ —in contradiction with Remark 7.12.
- (4) Luckily, the above proof only involves typable λY -terms. So we only need these properties to hold for all $M \in \Lambda_Y^\rightarrow$. We conjecture that property II indeed holds for all reducing fpc's Y and $M \in \Lambda_Y^\rightarrow$.
- (5) Our proof of property I for a class of reducing fpc's is obtained by considering a larger class of weak fpc's, namely the hereditarily reducing ones.

7.4 Hereditarily reducing weak fpc's

In order to state property I in its most general form, we now introduce the class of hereditarily reducing weak fpc's.

We have seen in Definition 3.5(ii) that an fpc Y is reducing whenever $Yx \rightarrow_\beta x(Yx)$. The problem is that the set of reducing fpc's is *not* closed under β -reduction, as shown by the following counterexample.

EXAMPLE 7.20

Let us consider the following variant Θ^I of Turing's fpc:

$$\Theta^I = \mathbf{W}^I \mathbf{W}^I, \quad \mathbf{W}^I = \lambda wx.x(\mathbf{I}wx).$$

It is easy to check that Θ^I is reducing. Obviously, $\mathbf{W}^I \rightarrow_\beta \mathbf{W}$ and hence $\Theta^I \rightarrow_\beta \Theta$, but if we only contract $\mathbf{I}w$ in the *second* occurrence of \mathbf{W}^I in Θ^I , we obtain the fpc

$$\Theta' = \mathbf{W}^I \mathbf{W},$$

which is no longer reducing.

This situation motivates the introduction of the following notion. It amounts to relaxing the requirement $Yx \rightarrow_\beta x(Yx)$ to mere syntactic separability of x from Y .

DEFINITION 7.21

A weak fpc $Y \in \Lambda$ is *hereditarily reducing* whenever it satisfies the following property:

$$\begin{aligned} \forall K \geq 0, \forall x \notin \text{FV}(Y), \forall N \in \Lambda \text{ such that } Yx \rightarrow_\beta N, \\ \exists k \geq K, \exists Y^* \in \Lambda \text{ such that } x \notin \text{FV}(Y^*) \text{ \& } N \rightarrow_\beta x^k(Y^*x). \end{aligned} \quad (7)$$

We denote by \mathcal{Y} the set of all hereditarily reducing weak fpc's.

While the above definition might seem quite intricate at first, its essential meaning is borne in the requirement that $x \notin \text{FV}(Y^*)$. Indeed, \mathcal{Y} consists of all weak fpc's Y such that any reduction starting with Yx can be continued until the variable x is *once again separated*, on the syntactic level, from the 'engine' producing the infinite BT $x(x(x(\dots)))$.

LEMMA 7.22

- (1) If Y is a terminal fpc then $Y \in \mathcal{Y}$.
- (2) If Y is a weak fpc, then $Y\delta \in \mathcal{Y}$.

PROOF. (1) It follows easily from the definition of terminal fpc's (see Definition 3.5).

(2) We divide the proof into claims.

CLAIM 11

Let $C[]$ be a context such that $\lambda x.C[x]$ is a weak fpc and $x \notin \text{FV}(C[])$. For every $n \in \mathbb{N}$ there exists a weak fpc Y' such that $C[\delta]x \rightarrow_\beta x^n(Y'x)$.

SUBPROOF. Proceed by induction on n . In case $n = 0$, we can simply take $Y' = C[\delta]$. Otherwise, for all $z \notin \text{FV}(C[])$ there is $N \in \Lambda$ such that $(\lambda x.C[x])z \rightarrow_\beta zN$ and since the latter is a weak hnf it can be reached by performing weak head reduction:

$$(\lambda x.C[x])z \rightarrow_w C[z] \rightarrow_w zN' \rightarrow_\beta zN$$

for some $N' \in \Lambda$. Notice that $\lambda z.N, \lambda z.N'$ must be weak fpc's as well. As weak head reductions are closed under substitution, we obtain (using $[\delta/z]$)

$$(\lambda x.C[x])\delta x \rightarrow_w C[\delta]x \rightarrow_w \delta D[\delta]x \rightarrow_w (\lambda x.xD[\delta]x)x \rightarrow_w x(D[\delta]x)$$

for some $D[]$ such that $D[\delta]$ is again a weak fpc, so we conclude by induction hypothesis. \square

CLAIM 12

Let $\lambda y.C[y]$ be a weak fpc and $x, y \notin \text{FV}(C[])$. For all reduction sequences $C[\delta]x \rightarrow_\beta N$ there exist $Z \in \Lambda, n \in \mathbb{N}$ such that $N \rightarrow_\beta x^n(Zx)$ and $x \notin \text{FV}(Z)$.

SUBPROOF. By induction on the length of the standard reduction $\rho : C[\delta]x \rightarrow_s N$, which exists by the standardization theorem for Λ . There are two cases.

Case 1: All reductions in ρ happen in the context $[], x$, in other words $C[\delta]$ reduces but does not 'eat' the x . In this case, N has already the correct form for $n = 0$ because $C[\delta]$ cannot create the variable x along its reduction. (This case includes the degenerate case of an empty reduction sequence.)

Case 2: Otherwise, the standard reduction ρ must have the following form

$$C[\delta]x \rightarrow_w \delta D[\delta]x \rightarrow_w (\lambda x.x(D[\delta]x))x \rightarrow_w x(D[\delta]x) \rightarrow_s N,$$

which entails $N = xN'$ for some N' satisfying $\sigma : D[\delta]x \rightarrow_s N'$ for a shorter (possibly empty) standard reduction σ . As $\lambda y.D[y]$ is a weak fpc we conclude by applying the induction hypothesis. \square

Now, since every weak fpc $Y \rightarrow_\beta \lambda x.C[x]$ for an appropriate context $C[\]$, Claim 11 entails that $Y\delta x \rightarrow_\beta x^k(Y'x)$ for arbitrarily large k . Therefore, $Y\delta \in \mathcal{Y}$ follows from Claim 12 by applying Church–Rosser.

The property that Y weak fpc entails $Y \in \mathcal{Y}$ is false, as evidenced by the following example.

EXAMPLE. Given a reducing fpc Y , consider $\mathcal{B} = \lambda x.Y(\mathbf{B}x)x$, which is a modified version of the Bible fpc. Clearly $\mathcal{B}x \rightarrow_\beta Y(\mathbf{B}x)x \rightarrow_\beta \mathbf{B}x(Y(\mathbf{B}x))x \rightarrow_\beta x(Y(\mathbf{B}x))x \rightarrow_\beta x^k(Y(\mathbf{B}x)x)$. Moreover, this reduction sequence is unavoidable in the construction of its BT; therefore, $\mathcal{B} \notin \mathcal{Y}$ since $x \in \text{FV}(Y(\mathbf{B}x))$.

DEFINITION 7.23

Let $Y, Y' \in \Lambda$, and let $x \notin \text{FV}(YY')$. Define

$$Y \Downarrow_k Y' \text{ if and only if } Yx \rightarrow_\beta x^k(Y'x).$$

We write $Y \Downarrow Y'$ whenever $Y \Downarrow_k Y'$ holds for some $k \in \mathbb{N}$.

PROPOSITION 7.24

Let $Y \in \mathcal{Y}$.

- (1) If $Y \Downarrow Y'$ then $Y' \in \mathcal{Y}$. Hence, \mathcal{Y} is closed under \Downarrow .
- (2) If $Y \rightarrow_\beta Y'$ then $Y \Downarrow Y'$. Hence, \mathcal{Y} is closed under β -reduction.
- (3) If $Y' \rightarrow_\beta Y$ then $Y' \in \mathcal{Y}$. Hence, \mathcal{Y} is closed under β -conversion.
- (4) For $k, k' \in \mathbb{N}$, $Y \Downarrow_k Y'$ and $Y' \Downarrow_{k'} Y''$ entail $Y \Downarrow_{k+k'} Y''$. So \Downarrow is transitive.

PROOF. Fix $Y \in \mathcal{Y}$.

- (1) Suppose that $Y \Downarrow Y'$ because, say, $Y \Downarrow_{k_0} Y'$.
Toward $Y' \in \mathcal{Y}$, let $K \in \mathbb{N}$, $Y'x \rightarrow_\beta Y''$ be given.
From $Yx \rightarrow_\beta x^{k_0}(Y'x)$, $Y'x \rightarrow_\beta Y''$, we get $Yx \rightarrow_\beta x^{k_0}(Y'')$.
Since $Y \in \mathcal{Y}$, let $k \geq K + k_0$, $Y^* \in \Lambda$, $x \notin \text{FV}(Y^*)$ be such that

$$x^{k_0}(Y'') \rightarrow_\beta x^k(Y^*x) = x^{k_0+K+k'}(Y^*x).$$

Since the above reduction is entirely in Y'' , $Y'' \rightarrow_\beta x^{K+k'}(Y^*x)$.

- (2) If $Y \rightarrow_\beta Y'$, then $Yx \rightarrow_\beta x^0(Y'x)$, so $Y \Downarrow_0 Y'$. By (1), $Y' \in \mathcal{Y}$.
- (3) Assume $Y' \rightarrow_\beta Y$. Then $Y'x \rightarrow_\beta Yx$.
Let $Y'x \rightarrow_\beta Z$, $K \geq 0$ be given.
By Church–Rosser, there exists a Z^0 such that $Yx \rightarrow_\beta Z^0 \beta \leftarrow Z$.
Since $Y \in \mathcal{Y}$, there exists a $k \geq K$, and a reduction $Z^0 \rightarrow_\beta x^k(Z'x)$.
So we have $Z \rightarrow_\beta Z^0 \rightarrow_\beta x^k(Z'x)$, as required.
- (4) From $Yx \rightarrow_\beta x^k(Y'x)$, $Y'x \rightarrow_\beta x^{k'}(Y''x)$, we immediately find that

$$Yx \rightarrow_\beta x^{k+k'}(Y''x).$$

\square

LEMMA 7.25

The set \mathcal{Y} contains all double fpc's.

PROOF. If Y is a double fpc, then $Y =_{\beta} Y\delta$. By Lemma 7.22(2), $Y\delta \in \mathcal{Y}$. By Proposition 7.24(3) \mathcal{Y} is closed under $=_{\beta}$, so we get $Y \in \mathcal{Y}$. \square

7.5 Non-uniform reduction extension properties

To allow for the fact that weak fpc's may change at various stages on the BT, the statements of the reduction extension properties need to be refined accordingly.

NOTATION 7.26

Let \rightarrow_r be a notion of reduction for λY , and let Y, Y' be weak fpc's. We write

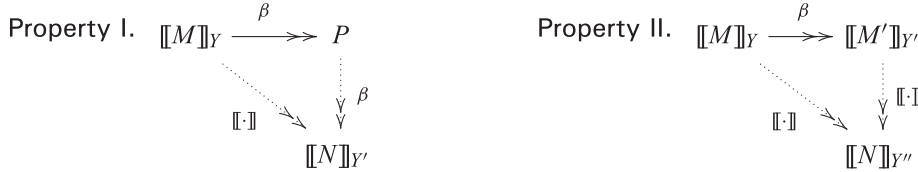
$$\llbracket M \rrbracket_Y \xrightarrow{\llbracket \cdot \rrbracket}_r \llbracket N \rrbracket_{Y'} \text{ if and only if } M \rightarrow_r N \text{ and } Y \not\Downarrow Y'.$$

As special cases, we consider

- $\llbracket M \rrbracket_Y \xrightarrow{\llbracket \cdot \rrbracket} \llbracket N \rrbracket_{Y'}$ if $M \rightarrow_{\beta_Y} N$ and $Y \not\Downarrow Y'$.
- $\llbracket M \rrbracket_Y \xrightarrow{\llbracket \cdot \rrbracket} \llbracket N \rrbracket_{Y'}$ if $M \Rightarrow N$ and $Y \not\Downarrow Y'$.

DEFINITION 7.27 (Non-uniform reduction extension properties).

A weak fpc $Y \in \Lambda$ satisfies the *non-uniform reduction extension properties* if the following hold.



We now show that non-uniform property I holds for all hereditarily reducing weak fpc's. From now on, and until the end of the section, we consider fixed $Y \in \mathcal{Y}$ and $M \in \Lambda_Y$.

LEMMA 7.28

If $Y \not\Downarrow Y'$ holds then there exists $N \in \Lambda_Y$ such that $M \rightarrow_Y N$ and $\llbracket M \rrbracket_Y \rightarrow_{\beta} \llbracket N \rrbracket_{Y'}$.

PROOF. We proceed by structural induction on M .

- $M = x$: In this case we just take $N = x$.
- $M = \lambda x.M_0$: By definition, we have $\llbracket M \rrbracket_Y = \lambda x.\llbracket M_0 \rrbracket_Y$. By induction hypothesis, there exists $N_0 \in \Lambda_Y$ such that $M_0 \rightarrow_Y N_0$ and $\llbracket M_0 \rrbracket_Y \rightarrow_{\beta} \llbracket N_0 \rrbracket_{Y'}$ hold. As a consequence, $\lambda x.M_0 \rightarrow_Y \lambda x.N_0$. Moreover, $\llbracket M \rrbracket_Y = \lambda x.\llbracket M_0 \rrbracket_Y \rightarrow_{\beta} \lambda x.\llbracket N_0 \rrbracket_{Y'} = \llbracket \lambda x.N_0 \rrbracket_{Y'}$.
- $M = M_1 M_2$: By definition, we have $\llbracket M \rrbracket_Y = \llbracket M_1 \rrbracket_Y \llbracket M_2 \rrbracket_Y$. By induction hypothesis, there exist $N_1, N_2 \in \Lambda_Y$ such that $M_i \rightarrow_Y N_i$ and $\llbracket M_i \rrbracket_Y \rightarrow_{\beta} \llbracket N_i \rrbracket_{Y'}$ for $i \in \{1, 2\}$. As a consequence, $M_1 M_2 \rightarrow_Y N_1 N_2$. Moreover, $\llbracket M \rrbracket_Y = \llbracket M_1 \rrbracket_Y \llbracket M_2 \rrbracket_Y \rightarrow_{\beta} \llbracket N_1 \rrbracket_{Y'} \llbracket N_2 \rrbracket_{Y'} = \llbracket N_1 N_2 \rrbracket_{Y'}$.
- $M = Y M_3$: By definition, we have $\llbracket M \rrbracket_Y = \llbracket Y M_3 \rrbracket_Y = Y \llbracket M_3 \rrbracket_Y$. By induction hypothesis, there exists $N_3 \in \Lambda_Y$ such that $M_3 \rightarrow_Y N_3$ and $\llbracket M_3 \rrbracket_Y \rightarrow_{\beta} \llbracket N_3 \rrbracket_{Y'}$. Since $Y \not\Downarrow Y'$ holds, there exists $k \in \mathbb{N}$ such that $Yx \rightarrow_{\beta} x^k(Y'x)$. Setting $N = N_3^k(YN_3)$, we

have

$$\begin{aligned}
 M &= \mathbb{Y}M_3 \rightarrow_{\mathbb{Y}} \mathbb{Y}N_3 \rightarrow_{\mathbb{Y}} N_3(\mathbb{Y}N_3) \rightarrow_{\mathbb{Y}} N_3(N_3(\mathbb{Y}N_3)) \rightarrow_{\mathbb{Y}} \dots \\
 &\rightarrow_{\mathbb{Y}} N_3^k(\mathbb{Y}N_3) = N \\
 \llbracket M \rrbracket_Y &= Y\llbracket M_3 \rrbracket_Y \rightarrow_{\beta} Y\llbracket N_3 \rrbracket_{Y'} \\
 &= Yx[\llbracket N_3 \rrbracket_{Y'}/x] \\
 &\rightarrow_{\beta} x^k(Y'x)[\llbracket N_3 \rrbracket_{Y'}/x] \\
 &= \llbracket N_3 \rrbracket_{Y'}^k(Y'\llbracket N_3 \rrbracket_{Y'}) = \llbracket N_3^k(\mathbb{Y}N_3) \rrbracket_{Y'} = \llbracket N \rrbracket_{Y'}.
 \end{aligned}$$

□

LEMMA 7.29

If $\llbracket M \rrbracket_Y \Rightarrow M'$ then there exist $N \in \Lambda_{\mathbb{Y}}$ and $Y' \in \mathcal{Y}$ such that $\llbracket M \rrbracket_Y \xrightarrow{\llbracket \cdot \rrbracket} \llbracket N \rrbracket_{Y'}$ and $M' \rightarrow_{\beta} \llbracket N \rrbracket_{Y'}$.

PROOF. We proceed by induction on a derivation of $\llbracket M \rrbracket_Y \Rightarrow M'$.

$\llbracket M \rrbracket_Y = x \Rightarrow x$: Here $M' = x$, so we can take $Y' = Y$ and $N = M = x$. Then certainly $M \Rightarrow N = x$, $M' \rightarrow_{\beta} \llbracket N \rrbracket_{Y'} = x$.

$\llbracket M \rrbracket_Y = \lambda x. \llbracket M_0 \rrbracket \Rightarrow \lambda x. M'_0$: Here $M = \lambda x. M_0$, $\llbracket M_0 \rrbracket \Rightarrow M'_0$ and $M' = \lambda x. M'_0$. By induction hypothesis, there are $N_0 \in \Lambda_{\mathbb{Y}}$ and $Y' \in \mathcal{Y}$ such that $Y \not\Downarrow Y'$ and $M_0 \Rightarrow N_0$ with $M'_0 \rightarrow_{\beta} \llbracket N_0 \rrbracket_{Y'}$. Letting $N = \lambda x. N_0$, we verify

$$\begin{aligned}
 M &= \lambda x. M_0 \Rightarrow \lambda x. N_0 = N \\
 M' &= \lambda x. M'_0 \rightarrow_{\beta} \lambda x. \llbracket N_0 \rrbracket_{Y'} = \llbracket N \rrbracket_{Y'}.
 \end{aligned}$$

$\llbracket M \rrbracket_Y = UV \Rightarrow M'_1 M'_2$: In this case, there are two possibilities:

- $M = \mathbb{Y}M_2$, $U = Y \Rightarrow M'_1$ and $V = \llbracket M_2 \rrbracket_Y \Rightarrow M'_2$.
By induction hypothesis, there are $N_2 \in \Lambda_{\mathbb{Y}}$ and $Y_2 \in \mathcal{Y}$ such that $Y \not\Downarrow Y_2$ and $M_2 \Rightarrow N_2$ with $M'_2 \rightarrow_{\beta} \llbracket N_2 \rrbracket_{Y_2}$.
Since $Y \Rightarrow M'_1$ and $Y \not\Downarrow Y_2$, we get

$$M'_1 x \beta \leftarrow Yx \rightarrow_{\beta} x^{k_2}(Y_2 x).$$

By Church–Rosser, there exist $k'_2 \geq k_2$ and $Y'_2 \in \Lambda$ such that

$$M'_1 x \rightarrow_{\beta} x^{k'_2}(Y'_2) \beta \leftarrow x^{k_2}(Y_2 x).$$

Now, using the fact that $Y \in \mathcal{Y}$, we obtain $Y^* \in \Lambda$, $k^* \geq k'_2$ such that

$$x^{k'_2}(Y'_2) \rightarrow_{\beta} x^{k^*}(Y^* x)$$

and certainly $Y \not\Downarrow Y^*$. Moreover, we have (i) $M'_1 x \rightarrow_{\beta} x^{k^*}(Y^* x)$; therefore, $M'_1 \not\Downarrow Y^*$ and (ii) $Y_2 x \rightarrow_{\beta} x^{k^* - k_2}(Y^* x)$, and hence $Y_2 \not\Downarrow Y^*$. By Lemma 7.28, there exists $N_2^* \in \Lambda_{\mathbb{Y}}$ such that

$$N_2 \rightarrow_{\mathbb{Y}} N_2^*, \quad \llbracket N_2 \rrbracket_{Y_2} \rightarrow_{\beta} \llbracket N_2^* \rrbracket_{Y^*}.$$

- Now $M \Rightarrow \Upsilon N_2 \rightarrow_Y \Upsilon N_2^*$, and so $M \Rightarrow \Upsilon N_2^*$. Moreover, $M' = M'_1 M'_2 \rightarrow_\beta M'_1 \llbracket N_2 \rrbracket_{Y_2} \rightarrow_\beta M'_1 \llbracket N_2^* \rrbracket_{Y^*} \rightarrow_\beta \llbracket N_2^* \rrbracket_{Y^*}^{k^*} (Y^* \llbracket N_2^* \rrbracket_{Y^*}) = \llbracket N_2^{*k^*} (\Upsilon N_2^*) \rrbracket_{Y^*}$.
- $M = M_1 M_2$, $U = \llbracket M_1 \rrbracket_Y \Rightarrow M'_1$, $V = \llbracket M_2 \rrbracket_Y \Rightarrow M'_2$ and $M' = M'_1 M'_2$. By induction hypothesis, for $i \in \{1, 2\}$, there are $N_i \in \Lambda_Y$ and $Y_i \in \mathcal{Y}$ such that $Y \not\Downarrow Y_i$ and $M_i \Rightarrow N_i$ with $M'_i \rightarrow_\beta \llbracket N_i \rrbracket_{Y_i}$.

From the first of these conditions $Y \not\Downarrow Y_1$, $Y \not\Downarrow Y_2$, we get

$$x^{k_1}(Y_1 x) \rightarrow_\beta \leftarrow_\beta Yx \rightarrow_\beta x^{k_2}(Y_2 x)$$

By Church–Rosser, there exist $k_{12} \geq \max\{k_1, k_2\}$ and $Y_{12} \in \Lambda$ such that

$$x^{k_1}(Y_1 x) \rightarrow_\beta x^{k_{12}}(Y_{12}) \rightarrow_\beta \leftarrow_\beta x^{k_2}(Y_2 x).$$

Now, using the fact that $Y \in \mathcal{Y}$, we obtain $Y^* \in \Lambda$, $k^* \geq k_{12}$ such that

$$x^{k_{12}}(Y_{12}) \rightarrow_\beta x^{k^*}(Y^* x)$$

and certainly $Y \not\Downarrow Y^*$. Moreover, for $i \in \{1, 2\}$, we have $Y_i x \rightarrow_\beta x^{k^* - k_i}(Y^* x)$ and therefore $Y_i \not\Downarrow Y^*$. By Lemma 7.28, there exist $N_1^*, N_2^* \in \Lambda_Y$ such that

$$N_i \rightarrow_Y N_i^*, \quad \llbracket N_i \rrbracket_{Y_i} \rightarrow_\beta \llbracket N_i^* \rrbracket_{Y^*}. \quad (i \in \{1, 2\})$$

Now $M = M_1 M_2 \Rightarrow N_1 N_2 \rightarrow_Y N_1^* N_2^*$, and so $M \Rightarrow N_1^* N_2^*$. Moreover, $M' = M'_1 M'_2 \rightarrow_\beta \llbracket N_1 \rrbracket_{Y_1} \llbracket N_2 \rrbracket_{Y_2} \rightarrow_\beta \llbracket N_1^* \rrbracket_{Y^*} \llbracket N_2^* \rrbracket_{Y^*} = \llbracket N_1^* N_2^* \rrbracket_{Y^*}$.

$\llbracket M \rrbracket_Y = (\lambda x. M_0) M_1 \Rightarrow M'_0[M'_1/x], M_i \Rightarrow M'_i$: In this case, there are two possibilities:

- $M = (\lambda x. P_0) P_1$ and $M_i = \llbracket P_i \rrbracket_Y$.
By induction hypothesis, we find Y_0, Y_1 with $Y \not\Downarrow Y_i$, $k_0, k_1 \geq 0$, and Q_0, Q_1 such that for each $i \in \{0, 1\}$ the following holds

$$Yx \rightarrow_\beta x^{k_i}(Y_i x),$$

$$P_i \Rightarrow Q_i,$$

$$M'_i \rightarrow_\beta \llbracket Q_i \rrbracket_{Y_i}.$$

As in the previous case, we first obtain Y^* with $Y_0, Y_1 \not\Downarrow Y^*$ such that

$$x^{k_0}(Y_0 x) \rightarrow_\beta x^{k^*}(Y^* x) \rightarrow_\beta \leftarrow_\beta x^{k_1}(Y_1 x).$$

Next, for each $i \in \{0, 1\}$, we apply Lemma 7.28 on $Y_i \not\Downarrow Y^*$ and Q_i in order to obtain Q'_0, Q'_1 satisfying

$$Q_i \rightarrow_Y Q'_i,$$

$$\llbracket Q_i \rrbracket_{Y_i} \rightarrow_\beta \llbracket Q'_i \rrbracket_{Y^*}.$$

From $P_i \Rightarrow Q_i \rightarrow_Y Q'_i$, we obtain $P_i \Rightarrow Q'_i$, and so

$$M = (\lambda x. P_0)P_1 \Rightarrow Q'_0[Q'_1/x].$$

Moreover, from $M'_i \rightarrow_\beta [Q_i]_{Y_i} \rightarrow_\beta [Q'_i]_{Y^*}$, we have $M'_i \rightarrow_\beta [Q'_i]_{Y^*}$ and therefore

$$M' = M'_0[M'_1/x] \rightarrow_\beta \llbracket Q'_0 \rrbracket_{Y^*}[\llbracket Q'_1 \rrbracket_{Y^*}/x] = \llbracket Q'_0[Q'_1/x] \rrbracket_{Y^*},$$

where the last equality is by the substitution lemma.

- $M = \Upsilon P, \lambda x. M_0 = Y$ and $M_1 = \llbracket P \rrbracket_Y$. Then $M_i \Rightarrow M'_i$ gives

$$Yx = (\lambda x. M_0)x \rightarrow_\beta M_0 \rightarrow_\beta M'_0$$

$$\llbracket P \rrbracket_Y \Rightarrow M'_1.$$

By induction hypothesis, we find $Y \Downarrow Y', P \Rightarrow Q$ with $M'_1 \rightarrow_\beta \llbracket Q \rrbracket_{Y'}$. Since $Y \Downarrow Y'$, there exists $k_0 \in \mathbb{N}$ such that $Yx \rightarrow_\beta x^{k_0}(Y'x)$. Notice that we also have $Yx \rightarrow_\beta M'_0$. Now let these reductions be joined

$$x^{k_0}(Y'x) \rightarrow_\beta Z \beta \leftarrow M'_0.$$

Using that $Y \in \mathcal{Y}$, let $k \geq k_0$, $Y^* \in \Lambda$ be such that $Z \rightarrow_\beta x^k(Y^*x)$. Then we obtain

$$x^{k_0}(Y'x) \rightarrow_\beta x^k(Y^*x) \beta \leftarrow M'_0.$$

In particular $Y' \Downarrow Y^*$, and by the previous lemma, there is $R \in \Lambda_Y$ such that

$$Q \rightarrow_Y R, \quad \llbracket Q \rrbracket_{Y'} \rightarrow_\beta \llbracket R \rrbracket_{Y^*}.$$

As $P \Rightarrow Q$ we get, setting $N = R^k(\Upsilon R)$, that

$$\Upsilon P \Rightarrow \Upsilon Q \rightarrow_Y \Upsilon R \rightarrow_Y R(\Upsilon R) \rightarrow_Y \cdots \rightarrow_Y R^k(\Upsilon R)$$

and so $M \Rightarrow N$. At the same time, from $M'_0 \rightarrow_\beta x^k(Y^*x)$, $M'_1 \rightarrow_\beta \llbracket Q \rrbracket_{Y'} \rightarrow_\beta \llbracket R \rrbracket_{Y^*}$ we find

$$\begin{aligned} M'_0[M'_1/x] &\rightarrow_\beta x^k(Y^*x)[\llbracket R \rrbracket_{Y^*}/x] \\ &= \llbracket R \rrbracket_{Y^*}^k(Y^* \llbracket R \rrbracket_{Y^*}) \\ &= \llbracket R^k(\Upsilon R) \rrbracket_{Y^*} \end{aligned}$$

and hence $M' \rightarrow_\beta \llbracket N \rrbracket_{Y^*}$. Indeed, we also have $Y \Downarrow Y' \Downarrow Y^*$ from which we conclude by transitivity. \square

We are now ready to prove that non-uniform reduction extension property I holds for all $Y \in \mathcal{Y}$.

THEOREM 7.30

Let $Y \in \mathcal{Y}$. For all $M \in \Lambda_Y, M' \in \Lambda$:

$$\llbracket M \rrbracket_Y \rightarrow_\beta M' \Rightarrow \exists(N, Y') : \llbracket M \rrbracket_Y \xrightarrow{\llbracket \cdot \rrbracket} \llbracket N \rrbracket_{Y'}, M' \rightarrow_\beta \llbracket N \rrbracket_{Y'}.$$

PROOF. By induction on the length of the reduction sequence $\llbracket M \rrbracket_Y \rightarrow_\beta M'$.

$\frac{\llbracket M \rrbracket_Y = M'}{\llbracket M \rrbracket_Y \rightarrow_\beta M_0 \rightarrow_\beta M'}:$ Take $N = M, Y' = Y$.
 By induction hypothesis, there are $N_0 \in \Lambda_Y, Y \Downarrow Y_0$ and reductions

$$\rho : M \rightarrow_{\beta_Y} N_0$$

$$\sigma : M_0 \rightarrow_\beta \llbracket N_0 \rrbracket_{Y_0}.$$

Projecting the redex R contracted in $M_0 \xrightarrow{R}_\beta M'$ over σ induces a parallel reduction

$$R \setminus \sigma : \llbracket N_0 \rrbracket_{Y_0} \Rightarrow M_1$$

and the finite reduction $\sigma \setminus R : M' \rightarrow_\beta M_1$.

$\llbracket M \rrbracket_Y \rightarrow_\beta M_0 \rightarrow_\beta M':$ By Lemma 7.29, we can find $N_1 \in \Lambda_Y, Y_0 \Downarrow Y_1$ and reductions

$$\rho_1 : N_0 \Rightarrow N_1$$

$$\sigma_1 : M_1 \rightarrow_\beta \llbracket N_1 \rrbracket_{Y_1}.$$

Now $\rho ; \rho_1 : M \rightarrow_{\beta_Y} N_0 \Rightarrow N_1$ clearly yields a reduction

$$M \rightarrow_{\beta_Y} N_1$$

and $\sigma \setminus R ; \sigma_1 : M' \rightarrow_\beta M_1 \rightarrow_\beta \llbracket N_1 \rrbracket_{Y_1}$ yields

$$M' \rightarrow_\beta \llbracket N_1 \rrbracket_{Y_1}.$$

Furthermore, $Y \Downarrow Y_0 \Downarrow Y_1$, from which we conclude since \Downarrow is transitive. \square

In the particular case of terminal (reducing) fpc's, the theorem above entails that also the reduction extension property I from Definition 7.17 holds.

COROLLARY 7.31

Every terminal fpc Y satisfies the reduction extension property I.

PROOF. If Y is terminal, then $Y \in \mathcal{Y}$ can be witnessed with $Y' = Y$ for any reduction starting from Yx . (That is, the fpc never changes.) In particular, the previous theorem is valid with $Y' = Y$.

That is, $\llbracket M \rrbracket_Y \rightarrow_\beta M'$ implies $\llbracket M \rrbracket_Y \xrightarrow{\llbracket \cdot \rrbracket_Y} \llbracket N \rrbracket_Y \beta \leftarrow M'$. \square

We end this section by presenting two conjectures; the first implies that non-uniform extension property I holds for all reducing fpc's (by Theorem 7.30), while the second entails the non-existence of double fixed point combinators in the simply typed setting.

CONJECTURE 1

If Y is a reducing fpc then $Y \in \mathcal{Y}$.

CONJECTURE 2

In the simply typed setting, every fpc $Y \in \mathcal{Y}$ satisfies non-uniform reduction extension property II for all $M \in \Lambda_Y^\rightarrow$.

Indeed, from Lemma 7.25 and Theorem 7.30 we get that every double fpc would satisfy non-uniform reduction property I. If Conjecture 2 holds, then Y moreover satisfies non-uniform reduction property II. Now, the same argument as in Proposition 7.18 applies; the interpretation of Λ_Y^\rightarrow reflects conversion, leading to the impossible λY -equality $\forall x = \forall \delta x$.

8 Conservativity of Double Fixed Point Operators

We analyse another possible proof technique, suggested by Klop, for proving the non-existence of double fixed point combinators. Consider the following λY -theory.

DEFINITION 8.1

Let δ^* be the λY -theory generated by the axiom $\forall x = Y\delta x$.

In [20], Klop raised the question of whether the λY -theory δ^* generated by the equation characterizing double fixed point combinators is a conservative extension of the λ -calculus. The motivation for this question is that, if this theory was found *not* to be conservative over Λ , this would immediately yield a proof of Statman's conjecture. Indeed, assuming that some fixed point combinator Y satisfies the equation $Y =_{\beta} Y\delta$, any equation between pure λ -terms that is provable with the axiom $\forall x = Y\delta x$ could be derived in the pure λ -calculus using Y , showing that δ^* is conservative over Λ .

The rest of the section is devoted to proving that Klop's question has a positive answer. This result shows that, unfortunately, this strategy cannot be used to settle Statman's conjecture.

8.1 The ν -reduction

To characterize equality in δ^* using standard rewriting techniques, we introduce a new notion of reduction:

$$Y\delta M \rightarrow YM. \quad \nu$$

LEMMA 8.2

For all $M, N \in \Lambda_Y$, $\delta^* \vdash M = N$ if and only if $M =_{\beta Y \nu} N$.

PROOF. (\Rightarrow) By definition, $=_{\beta Y \nu}$ is a contextual equivalence and therefore a λY -theory. A simple inspection of the ν -rule shows that $=_{\beta Y \nu}$ validates every axiom of the theory δ^* . We conclude since δ^* is the least λY -theory validating these axioms.

(\Leftarrow) This implication follows by an easy induction on the length of the conversion sequence $M = M_1 \leftrightarrow_{\beta Y \nu} \dots \leftrightarrow_{\beta Y \nu} M_k = N$. \square

The conservativity of δ^* will follow from the confluence property enjoyed by $\beta Y \nu$ -reduction. Note that this system is not (weakly) orthogonal, due to the overlap between λY - and ν -redexes. It is not terminating either; thus, Newman's lemma does not apply. Therefore, we need to prove confluence directly. As a first step, we show that ν -reduction enjoys the strong diamond property.

PROPOSITION 8.3

Let $M, N, P \in \Lambda_Y$. If $N \leftarrow_{\nu} M \rightarrow_{\nu} P$, then there exists $Q \in \Lambda_Y$ such that $N \rightarrow_{\nu} Q \leftarrow_{\nu} P$.

PROOF. Assume that $N \leftarrow_{\nu} M \rightarrow_{\nu} P$ by contracting the redexes L and R , respectively. If the two redexes are disjoint, then we easily close the diagram

$$N \xrightarrow{R/L}_{\nu} Q \xleftarrow{L/R}_{\nu} P.$$

Otherwise, one redex is contained in the other one, say, R occurs within L . Since L is an ν -redex it must have the shape $Y\delta N'$, so the occurrence of R must be contained in N' , witnessed by $N' \rightarrow_{\nu} P'$. That is, for some λY -context $C[\]$, we must have

$$N = C[YN'] \xleftarrow{\nu} C[Y\delta N'] = M \xrightarrow{R}_{\nu} C[Y\delta P'] = P.$$

We conclude since $C[\Upsilon N'] \xrightarrow{R/L}_v C[\Upsilon P'] \xleftarrow{L/R}_v C[\Upsilon \delta P']$. \square

The rest of the section is devoted to proving the confluence of $\beta\Upsilon v$ -reduction. We start by defining the parallel version of v -reduction and by studying its properties.

DEFINITION 8.4

The notion of *parallel v -reduction* \Rightarrow_v is given as the $\lambda\Upsilon$ -contextual closure of the following rule:

$$\frac{M \Rightarrow_v M'}{\Upsilon \delta M \Rightarrow_v \Upsilon M'}$$

PROPOSITION 8.5 (Postponement of v -reduction).

$$\begin{array}{ccc} 1. & \begin{array}{ccc} M & \xRightarrow{\quad} & M' \\ \downarrow \beta\Upsilon & & \downarrow \beta\Upsilon \\ N & \xRightarrow{\quad}_v & N' \end{array} & 2. & \begin{array}{ccc} M & \xRightarrow{\quad} & M' \\ \downarrow \beta\Upsilon & & \downarrow \beta\Upsilon \\ N & \xRightarrow{\quad}_v & N' \end{array} & 3. & \begin{array}{ccc} M & \xRightarrow{\quad} & M' \\ \downarrow \beta\Upsilon & & \downarrow \beta\Upsilon \\ N & \xRightarrow{\quad}_v & N' \end{array} \end{array}$$

PROOF. 1. We proceed by induction on the derivation of $M \Rightarrow_v M'$.

Case $x \Rightarrow_v x$. This case is impossible, because x has no $\beta\Upsilon$ -redex.

Case $\lambda x.M_0 \Rightarrow_v \lambda x.M'_0$ with $M_0 \Rightarrow_v M'_0$. Clearly, the redex contracted in $\lambda x.M'_0 \rightarrow_{\beta\Upsilon} N'$ must occur inside M'_0 , so that $N' = \lambda x.N'_0$ and $M'_0 \rightarrow_{\beta\Upsilon} N'_0$. By induction hypothesis, there exists a term N_0 such that $M_0 \rightarrow_{\beta\Upsilon} N_0 \Rightarrow_v N'_0$. Since reductions are contextual, we get $M = \lambda x.M_0 \rightarrow_{\beta\Upsilon} \lambda x.N_0 \Rightarrow_v \lambda x.N'_0 = N'$.

Case $M_1 M_2 \Rightarrow_v M'_1 M'_2$ with $M_i \Rightarrow_v M'_i$. We need to consider two subcases.

- If the redex contracted in $M'_1 M'_2 \rightarrow_{\beta\Upsilon} N'$ occurs inside some M'_i , say, in M'_1 , then by induction hypothesis we obtain that $M_1 \rightarrow_{\beta\Upsilon} N_1 \Rightarrow_v N'_1$, where $N' = N'_1 M'_2$. So we take $N = N_1 M_2$, and find $M = M_1 M_2 \rightarrow_{\beta\Upsilon} N_1 M_2 \Rightarrow_v N'_1 M'_2 = N'$.
- Otherwise, the redex occurs at the root in $M'_1 M'_2$. Since Υ cannot occur as a term on its own, the redex must be a β -redex. That is, $M'_1 = \lambda x.M'_{10}$, and $(\lambda x.M'_{10})M'_2 \rightarrow_{\beta} M'_{10}[M'_2/x] = N'$. In this case $M_1 \Rightarrow_v M'_1$ can only arise as $\lambda x.M_{10} \Rightarrow_v \lambda x.M'_{10}$, where $M_0 = \lambda x.M_{10}$ and $M_{10} \Rightarrow_v M'_{10}$. Therefore, we have $M = (\lambda x.M_{10})M_2 \rightarrow_{\beta} M_{10}[M_2/x] \Rightarrow_v M'_{10}[M'_2/x] = N'$.

Case $\Upsilon M_3 \Rightarrow_v \Upsilon M'_3$ with $M_3 \Rightarrow_v M'_3$. There are two subcases.

- If the redex contracted in $\Upsilon M'_3 \rightarrow_{\beta\Upsilon} N'$ occurs inside M'_3 , so that $N' = \Upsilon N'_3$ with $M'_3 \rightarrow_{\beta\Upsilon} N'_3$, then by induction hypothesis we have that $M_3 \rightarrow_{\beta\Upsilon} N_3 \Rightarrow_v N'_3$ and hence that $M = \Upsilon M_3 \rightarrow_{\beta\Upsilon} \Upsilon N_3 \Rightarrow_v \Upsilon N'_3 = N'$.
- Otherwise, the redex contracted in $\Upsilon M'_3 \rightarrow_{\beta\Upsilon} N'$ is the Υ -redex at the root, and its contractum N' is $M'_3(\Upsilon M'_3)$. From $M_3 \Rightarrow_v M'_3$, we get $\Upsilon M_3 \Rightarrow_v \Upsilon M'_3$, which entails $M_3(\Upsilon M_3) \Rightarrow_v M'_3(\Upsilon M'_3)$. Therefore, $M = \Upsilon M_3 \rightarrow_{\beta\Upsilon} M_3(\Upsilon M_3) \Rightarrow_v M'_3(\Upsilon M'_3) = N'$.

Case $\Upsilon \delta M_4 \Rightarrow_v \Upsilon M'_4$ with $M_4 \Rightarrow_v M'_4$. Again, there are two subcases.

- If the redex contracted in $\Upsilon M'_4 \rightarrow_{\beta\Upsilon} N'$ occurs inside M'_4 , with $M'_4 \rightarrow_{\beta\Upsilon} N'_4$ and $N' = \Upsilon N'_4$, then induction hypothesis yields $M_4 \rightarrow_{\beta\Upsilon} N_4 \Rightarrow_v N'_4$. From this it follows that $M = \Upsilon \delta M_4 \rightarrow_{\beta\Upsilon} \Upsilon \delta N_4 = N$ and $N \Rightarrow_v \Upsilon N'_4 = N'$.

- Otherwise, the redex contracted in $\Upsilon M'_4 \rightarrow_{\beta\Upsilon} N'$ is the root redex, and N' is its contractum $M'_4(\Upsilon M'_4)$. We have

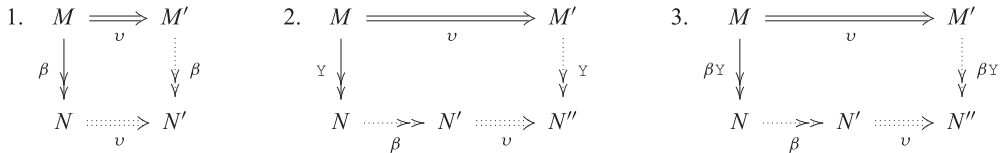
$$\begin{aligned} M &= \Upsilon \delta M_4 \rightarrow_{\Upsilon} \delta(\Upsilon \delta M_4) \rightarrow_{\beta} (\lambda x.x(\Upsilon \delta x))M_4 \\ &\rightarrow_{\beta} M_4(\Upsilon \delta M_4) \\ &\Rightarrow_v M'_4(\Upsilon M'_4) = N', \end{aligned}$$

where the \Rightarrow_v -step arises by combining $M_4 \Rightarrow_v M'_4$ with $\Upsilon \delta M_4 \Rightarrow_v \Upsilon M'_4$ using the application rule.

2. By induction on $M' \rightarrow_{\beta\Upsilon} N'$, tiling 1 vertically.
3. By induction on $M \rightarrow_v M'$, tiling 2 horizontally. \square

Remark. Notice that the above proof can be refined to a postponement of v reduction along β reduction instead of $\beta\Upsilon$ reduction.

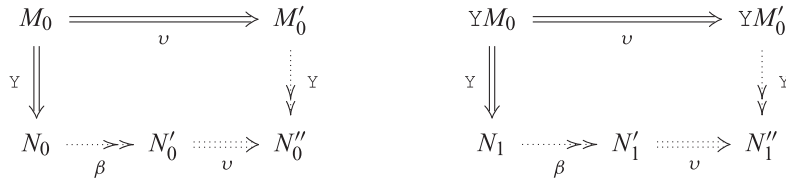
LEMMA 8.6 (Commutations of v -reductions).



PROOF. 1. This is immediate since the rules for β - and v -reductions are orthogonal.

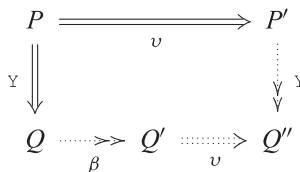
2. Recall that \Rightarrow denotes the parallel $\beta\Upsilon$ -reduction introduced in Definition 7.4. Since, by Corollary A5, $M \rightarrow_{\Upsilon} N$ entails $M \Rightarrow N$ we proceed by induction on the derivation of the latter. The only interesting cases arise when M is a Υ -redex or an v -redex.

Case $\Upsilon M_0 \Rightarrow N_0 N_1$ with $M_0 \Rightarrow N_0$ and $\Upsilon M_0 \Rightarrow N_1$. As $\Upsilon M_0 \Rightarrow_v M' = \Upsilon M'_0$, we must have $M_0 \Rightarrow_v M'_0$. By induction hypothesis, we can complete the diagram



From $M'_0 \rightarrow_{\Upsilon} N''_0$ and $\Upsilon M'_0 \rightarrow_{\Upsilon} N'_1$ we get $\Upsilon M'_0 \rightarrow_{\Upsilon} N''_0 N'_1$. So we have $N = N_0 N_1 \rightarrow_{\beta} N'_0 N'_1 \Rightarrow_v N''_0 N'_1 \Upsilon \leftarrow \Upsilon M'_0 = M'$.

Case $\Upsilon \delta P \Rightarrow \Upsilon \delta Q$ with $P \Rightarrow Q$. Suppose moreover that $\Upsilon \delta P \Rightarrow_v \Upsilon P'$ with $P \Rightarrow_v P'$. By induction hypothesis, we can complete the diagram



872 The fixed point property and a technique to harness double fixed point combinators

Thus, we have $N = \Upsilon\delta Q \rightarrow_{\beta} \Upsilon\delta Q' \Rightarrow_{\nu} \Upsilon Q'' \Upsilon \leftarrow \Upsilon P' = M'$.

Case $\Upsilon\delta P \Rightarrow N_1 N_2 Q$ with $\delta \Rightarrow N_1, \Upsilon\delta \Rightarrow N_2$ and $P \Rightarrow Q$. We also suppose that $\Upsilon\delta P \Rightarrow_{\nu} \Upsilon P'$ with $P \Rightarrow_{\nu} P'$. The fact that δ is a normal form entails $N_1 = \delta$, so we obtain $\Upsilon\delta P \Rightarrow \delta N_2 Q$. Since the only Υ -reducts of $\Upsilon\delta$ are λY -terms of the form $\delta^k(\Upsilon\delta)$ for some k we must have $N_2 = \delta^k(\Upsilon\delta)$. We also have the β -reduction:

$$\begin{aligned} N = N_1 N_2 Q &= \delta(\delta^k(\Upsilon\delta))Q \rightarrow_{\beta}^2 Q(\delta^k(\Upsilon\delta)Q) \\ &= Q(\delta(\delta^{k-1}(\Upsilon\delta))Q) \rightarrow_{\beta}^2 Q(Q(\delta^{k-1}(\Upsilon\delta)Q)) \\ &\quad \rightarrow_{\beta}^{2(k-1)} Q^{k+1}(\Upsilon\delta Q). \end{aligned}$$

By induction hypothesis, we have $Q \rightarrow_{\beta} R \Rightarrow_{\nu} S \Upsilon \leftarrow P'$. Therefore,

$$\begin{array}{ccc} \Upsilon\delta P & \xRightarrow{\quad \quad \quad \nu \quad \quad \quad} & \Upsilon P' \\ \Upsilon \downarrow & & \downarrow \Upsilon \\ \delta(\delta^k(\Upsilon\delta))Q & \xrightarrow[\beta]{\quad \quad \quad} Q^{k+1}(\Upsilon\delta Q) \xrightarrow[\beta]{\quad \quad \quad} R^{k+1}(\Upsilon\delta R) \xrightarrow[\nu]{\quad \quad \quad} S^{k+1}(\Upsilon S) \end{array}$$

Notice that $R^{k+1}(\Upsilon\delta R) \Rightarrow_{\nu} S^{k+1}(\Upsilon S)$ is obtained by putting together the reduction $R \Rightarrow_{\nu} S$ at $k+2$ disjoint positions, while using a single ν -reduction step to remove the δ occurring at depth $k+1$. The Υ -reduction $\Upsilon P' \rightarrow_{\Upsilon} S^{k+1}(\Upsilon S)$ is obtained as

$$\begin{aligned} \Upsilon P' &\rightarrow_{\Upsilon} P'(\Upsilon P') \rightarrow_{\Upsilon} S(\Upsilon P') \rightarrow S(P'(\Upsilon P')) \rightarrow_{\Upsilon} S(S(\Upsilon P')) \rightarrow_{\Upsilon} \dots \\ &\rightarrow_{\Upsilon} S^{k+1}(\Upsilon P') \rightarrow_{\Upsilon} S^{k+1}(\Upsilon S). \end{aligned}$$

3. We proceed by induction on the length n of the reduction $M \rightarrow_{\beta\Upsilon} N$.

Case $n = 0$. In this case $M = N$ and there is nothing to prove.

Case $n > 0$. The reduction $M \rightarrow_{\beta\Upsilon} N$ factors as $M \rightarrow_{\beta\Upsilon} N_0 \rightarrow_{\beta\Upsilon} N$. By applying the induction hypothesis to $M \rightarrow_{\beta\Upsilon} N_0$ of length $n-1$, we have

$$\begin{array}{ccc} M & \xRightarrow{\quad \quad \quad \nu \quad \quad \quad} & M' \\ \beta\Upsilon \downarrow & & \downarrow \beta\Upsilon \\ N_0 & \xrightarrow[\beta]{\quad \quad \quad} N'_0 \xrightarrow[\nu]{\quad \quad \quad} N''_0 \end{array}$$

There are two subcases.

- If $N_0 \rightarrow_{\beta\Upsilon} N$ contracts a β -redex, then we conclude by

$$\begin{array}{ccccc} N_0 & \xrightarrow[\beta]{\quad \quad \quad} & N'_0 & \xRightarrow{\quad \quad \quad \nu \quad \quad \quad} & N''_0 \\ \beta \downarrow & & \downarrow \beta & & \downarrow \beta \\ N & \xrightarrow[\beta]{\quad \quad \quad} & N' & \xrightarrow[\nu]{\quad \quad \quad} & N'' \end{array}$$

where the square on the left exists by confluence of β , and the one on the right by part 1.

- If $N_0 \rightarrow_{\beta\gamma} N$ contracts a γ -redex, then we are done since

$$\begin{array}{ccccc}
 N_0 & \xrightarrow{\beta} & N'_0 & \xrightarrow{v} & N''_0 \\
 \gamma \downarrow & & \downarrow \gamma & & \downarrow \gamma \\
 N & \xrightarrow{\beta} & N' & \xrightarrow{\beta} & N'' & \xrightarrow{v} & N'''
 \end{array}$$

where the square on the left exists by commutation of β and γ (which holds by orthogonality), and the one on the right by part 2. \square

PROPOSITION 8.7

$$\begin{array}{ccc}
 1. & M \xrightarrow{v} M' & 2. & M \xrightarrow{\beta\gamma v} M' \\
 \beta\gamma \downarrow & & \beta\gamma v \downarrow & & \beta\gamma v \downarrow \\
 N \xrightarrow{\beta} N' & \xrightarrow{v} N'' & N \xrightarrow{\beta\gamma v} N' & & N \xrightarrow{\beta\gamma v} N'
 \end{array}$$

PROOF. 1. We proceed by induction on the length of $M \rightarrow_v M'$, omitting the base case, which is trivial. The inductive case is obtained via the following diagram:

$$\begin{array}{ccccccc}
 M & \xrightarrow{v} & M_1 & \xrightarrow{v} & M' \\
 \beta\gamma \downarrow & & \downarrow \beta\gamma & & \downarrow \beta\gamma \\
 N & \xrightarrow{\beta} & N'_0 & \xrightarrow{v} & N_1 & \xrightarrow{\beta} & N'_k & \xrightarrow{v} & N''
 \end{array}$$

by Lemma 7.6(3) by IH

By applying postponement of v -reduction to the bottom row, one turns $N'_0 \rightarrow_v N_1 \rightarrow_{\beta} N'_k$ into $N_0 \rightarrow_{\beta} N' \rightarrow_v N'_k$ for some N' from which the result follows immediately.

2. Given the reductions $M \rightarrow_{\beta\gamma v} N$ and $M \rightarrow_{\beta\gamma v} M'$, we first apply v -postponement to each and then we complete the diagram

$$\begin{array}{ccccc}
 M & \xrightarrow{\beta\gamma} & \cdot & \xrightarrow{v} & M' \\
 \downarrow \beta\gamma & & \downarrow \beta\gamma & & \downarrow \beta\gamma \\
 \cdot & \xrightarrow{\beta\gamma} & \cdot & \xrightarrow{\beta} & \cdot & \xrightarrow{v} & \cdot \\
 \downarrow v & & \downarrow v & & \downarrow v & & \downarrow v \\
 N & \xrightarrow{\beta\gamma} & \cdot & \xrightarrow{\beta} & \cdot & \xrightarrow{v} & N'
 \end{array}$$

The squares that appear along the main diagonal are obtained by confluence of β_Y -, β - and ν -reductions, individually. The rectangles covering the bottom-left and top-right corners are given by part 1. The remaining squares follow by commutation of ν - and β -reductions (Lemma 8.6(1)). \square

By the well-known Theorem 3.1.12 in [2], we obtain the following corollary.

COROLLARY 8.8

Let $M, N \in \Lambda_Y$. If $M =_{\beta_Y \nu} N$ then there exists a λ_Y -term Z such that $M \rightarrow_{\beta_Y \nu} Z \beta_{Y \nu} \leftarrow N$.

THEOREM 8.9

The λ_Y -theory δ^* is a conservative extension of λ -calculus.

PROOF. Let $M, N \in \Lambda$ and suppose that $\delta^* \vdash M = N$. By Lemma 8.2 we have $M =_{\beta_Y \nu} N$ and, by Corollary 8.8, there exists a λ_Y -term Z such that

$$M \rightarrow_{\beta_Y \nu} Z \beta_{Y \nu} \leftarrow N.$$

Since none of the reduction rules are able to create a new occurrence of the symbol Y , there is no point in these reductions where Y - or ν -redexes can appear. Thus, the reductions above are actually β -reductions, so we conclude that $M =_{\beta} N$. \square

9 Conclusions

We have investigated two questions concerning (sets of) fixed points of terms in λ -calculus, the veracity of the *fixed point property*, and the *existence of a double fixed point combinator*. We have provided partial answers to both questions, and established several promising new techniques for tackling full solutions.

One novel aspect of the present work is to consider the questions in different λ -theories. For example, we have devised an example showing that the fixed point property patently fails in any sensible lambda theory, thus proving a conjecture of Intrigila and Statman.

Apart from the major problem of settling the status of the two main questions in the most fine-grained λ -theory—i.e. the ‘usual’ theory whose equivalence classes consist of terms that are β -equivalent—several lesser open problems remain; e.g. providing a characterization of the fixed point property in semi-sensible theories, and investigating the usefulness of the novel technique for refuting the existence of double fixed points combinators in the setting of (simple) types. We urge the reader to peruse the conjectures and suggestions that occur throughout the paper, both explicitly and in the running text.

A Technical Appendix

A.1 Standardization

The standard reduction can be thought of as a ‘canonical serialization’ of the usual multistep reduction. This idea is made precise by the standardization theorem, which we now prove.

LEMMA A1 (Substitution lemma).

For $M, M', N, N' \in \Lambda_Y$, we have

- (i) $N \rightarrow_s N'$ implies $M[N/x] \rightarrow_s M[N'/x]$,
- (ii) $M \rightarrow_w M'$ and $N \rightarrow_s N'$ imply $M[N/x] \rightarrow_s M'[N'/x]$,
- (iii) $M \rightarrow_s M'$ and $N \rightarrow_s N'$ imply $M[N/x] \rightarrow_s M'[N'/x]$.

PROOF OF LEMMA A.1

- (i) By the fact that \rightarrow_s is a congruence.
- (ii) First notice that $M[N/x] \rightarrow_w M'[N/x]$. This can be seen by considering the possible shape of $M \rightarrow_w M'$, where

$$\begin{aligned}
 ((\lambda y.M_0)N_0 \cdots N_k)[N/x] &= (\lambda y.M_0[N/x])N_0[N/x] \cdots N_k[N/x] \\
 &\rightarrow_w M_0[N/x][N_0[N/x]/y]N_1[N/x] \cdots N_k[N/x] \\
 &= M_0[N_0/y][N/x]N_1[N/x] \cdots N_k[N/x] \\
 &= (M_0[N_0/y]N_1 \cdots N_k)[N/x] \\
 (\exists N_0 \cdots N_k)[N/x] &= \exists N_0[N/x] \cdots N_k[N/x] \\
 &\rightarrow_w N_0[N/x](\exists N_0[N/x])N_1[N/x] \cdots N_k[N/x] \\
 &= (N_0(\exists N_0)N_1 \cdots N_k)[N/x].
 \end{aligned}$$

Next, we have $M'[N/x] \rightarrow_s M'[N'/x]$ by point (i).

- (ii) Thus, $M[N/x] \rightarrow_w M'[N/x] \rightarrow_s M'[N'/x]$.
- (ii) By the redex rule for \rightarrow_s , we have (ii).
- (iii) By induction on $M \rightarrow_s M'$, using (ii) in case of the redex rule. □

LEMMA A2

For $M, N \in \Lambda_Y$, we have that $M \rightarrow_s \lambda x.N$ entails $M \rightarrow_w \lambda x.M'$ and $M' \rightarrow_s N$ for some $M' \in \Lambda_Y$.

PROOF OF LEMMA A2. We proceed by induction on the derivation of $M \rightarrow_s \lambda x.N$. Since M reduces to an abstraction, there are only two possibilities:

- $M \rightarrow_s \lambda x.N$ because $M = \lambda x.M'$ and $M' \rightarrow_s N$. This case is trivial as $M \rightarrow_w \lambda x.M'$ follows from the reflexivity of \rightarrow_w .
- $M \rightarrow_s \lambda x.N$ because $M \rightarrow_w M_1$ and $M_1 \rightarrow_s \lambda x.N$. By induction hypothesis, there exists $M' \in \Lambda_Y$ such that $M_1 \rightarrow_w \lambda x.M'$ with $M' \rightarrow_s N$. Since $\rightarrow_w \subseteq \rightarrow_w$ and \rightarrow_w is transitive, we conclude $M \rightarrow_w \lambda x.M'$. □

THEOREM A3 (Standardization).

For all $M, N, N' \in \Lambda_Y$, we have

- (i) $M \rightarrow_s N \rightarrow_{\beta_Y} N'$ implies $M \rightarrow_s N'$,
- (ii) $M \rightarrow_s N \rightarrow_{\beta_Y} N'$ implies $M \rightarrow_s N'$,
- (iii) $M \rightarrow_{\beta_Y} N$ implies $M \rightarrow_s N$.

PROOF OF THEOREM A3. (i) By induction on the derivation of $M \rightarrow_s N$.

$M \rightarrow_w M' \rightarrow_s N \rightarrow_{\beta_Y} N'$: By induction hypothesis, we have $M' \rightarrow_s N'$. Now $M \rightarrow_w M' \rightarrow_s N'$, whence $M \rightarrow_s N'$.

$M = x \rightarrow_s x = N$: This case is inconsistent with $N \rightarrow_{\beta_Y} N'$.

$M = \lambda x.M_0 \rightarrow_s \lambda x.M'_0 = N$: Since $N = \lambda x.M'_0$ is not a redex, the redex contracted in $N \rightarrow_{\beta_Y} N'$ must occur below, in M'_0 . So $M'_0 \rightarrow_{\beta_Y} M''_0$ and $N' = \lambda x.M''_0$. By induction hypothesis, $M_0 \rightarrow_s M''_0$, whence $M = \lambda x.M_0 \rightarrow_s \lambda x.M''_0 = N'$.

$M = M_1 M_2, M_i \rightarrow_s M'_i$: We distinguish two subcases.

- The redex contracted in $N = M'_1 M'_2 \rightarrow_{\beta_Y} N'$ occurs at the root. Note that it cannot be a Y -redex, since Y cannot occur on its own. This entails that $M'_1 = \lambda x.M'_0$ is an abstraction and $N' = M'_0[M'_2/x]$, being the contractum of $(\lambda x.M'_0)M'_2$. Since $M_1 \rightarrow_s \lambda x.M'_0$, we get by Lemma A.2 a λY -term M_0 such that $M_1 \rightarrow_w \lambda x.M_0$ and $M_0 \rightarrow_s M'_0$. Therefore, we obtain $M = M_1 M_2 \rightarrow_w (\lambda x.M_0)M_2 \rightarrow_w M_0[M_2/x]$ on the one side. On the other side, we have $M_0 \rightarrow_s M'_0$ and $M_2 \rightarrow_s M'_2$. By the substitution lemma for standard reductions, we get a standard reduction $M_0[M_2/x] \rightarrow_s M'_0[M'_2/x]$. By an iterated application of the rule combining \rightarrow_w and \rightarrow_s to get a standard reduction, we obtain $M \rightarrow_s M'_0[M'_2/x] = N'$.
- The redex contracted in $N = M'_1 M'_2 \rightarrow_{\beta_Y} N'$ occurs below, in some M_i . So $M'_i \rightarrow_{\beta_Y} M''_i$, and $N' = M'_1 M'_2$, where we set $M'_{3-i} = M'_{3-i}$. By induction hypothesis, $M_1 \rightarrow_s M'_1$ and $M_2 \rightarrow_s M'_2$. Thus, $M = M_1 M_2 \rightarrow_s M'_1 M'_2 = N'$.

$M = YM_3 \rightarrow_s YM'_3 = N$: Again, we have two possibilities.

- The redex contracted in $N = YM'_3 \rightarrow_{\beta_Y} N'$ is the root redex. Then we have $N' = M'_3(YM'_3)$. From $M_3 \rightarrow_s M'_3$, we obtain $M_3(YM_3) \rightarrow_s M'_3(YM'_3)$. Now $M = YM_3 \rightarrow_w M_3(YM_3) \rightarrow_s M'_3(YM'_3) = N'$, whence $M \rightarrow_s N'$.
- The redex contracted in $YM'_3 \rightarrow_{\beta_Y} N'$ occurs in M'_3 . Then $N = YM'_3$, with $M'_3 \rightarrow_{\beta_Y} M''_3$. By induction hypothesis, we have that $M_3 \rightarrow_s M'_3$ holds. Now $M = YM_3 \rightarrow_s YM'_3 = N'$, which concludes the proof.

(ii) By straightforward induction on $N \rightarrow_{\beta_Y} N'$, using (i).

(iii) Immediate by (ii). □

A.2 Properties of parallel reduction

Notice that our definition of parallel reduction allows superdevelopment of newly created Y -redexes. While not strictly necessary, this simplifies some of our arguments.

One consequence of this is the following absorption lemma.

LEMMA A.4

The following rule is admissible:

$$\frac{M \Rightarrow_p M' \quad M' \rightarrow_Y N'}{M \Rightarrow_p N'}$$

PROOF OF LEMMA A.4. First, consider the length of the reduction $M' \rightarrow_Y N'$. When $M' \rightarrow_Y N'$ is empty, then $M' = N'$ and certainly $M \Rightarrow N'$. Otherwise, $M' \rightarrow_Y N \rightarrow_Y N'$, and induction yields that $M \Rightarrow N$. We now use a subsidiary induction on the derivation of this fact. Let Δ be the Y -redex contracted in the step $N \rightarrow_Y N'$.

$x \Rightarrow x = N$: This case is inconsistent with $N \xrightarrow{\Delta}_Y N'$.
 $M_1 M_2 \Rightarrow N_1 N_2 = N$: We are in a case where N is an application $N_1 N_2$ and therefore redex Δ fired in $N_1 N_2 \xrightarrow{\Delta} N'$ may not occur at the root because Y is not itself a term. So $\Delta \subseteq N_1$ or $\Delta \subseteq N_2$. Writing $N' = N'_1 N'_2$, we have $N_i \rightarrow_Y^{\leq 1} N'_i$. Since $M_i \Rightarrow N_i$, we have by induction hypothesis, that $M_i \Rightarrow N'_i$, whence $M \Rightarrow N'$.

$\lambda x.M_0 \Rightarrow \lambda x.N_0 = N$: Clearly, $\Delta \subseteq N_0$, so that $N_0 \rightarrow_Y P$ and $N' = \lambda x.P$. By induction, $M_0 \Rightarrow N_0 \rightarrow_Y P$ yields $M_0 \Rightarrow P$, and hence

$$M = \lambda x.M_0 \Rightarrow \lambda x.P = N'.$$

$(\lambda x.P)Q \Rightarrow P'[Q'/x] = N$: with $P \Rightarrow P'$ and $Q \Rightarrow Q'$. Since Y cannot occur as a term on its own, a Y -redex cannot be created by a substitution instance $P'[Q'/x]$. So Δ is inside either P' or Q' . That is, either $N' = P''[Q'/x]$, where $P' \xrightarrow{\Delta}_Y P''$, or $N' = P'[Q''/x]$, where $Q' \xrightarrow{\Delta}_Y Q''$. In either case, we can use induction hypothesis to get $P \Rightarrow P''$, respectively $Q \Rightarrow Q''$, and therefore $M \Rightarrow N'$.

$YP \Rightarrow YP' = N$: with $P \Rightarrow P'$. We split into two subcases.

- If $\Delta \subseteq P'$, so that $YP' \xrightarrow{\Delta}_Y YQ = N'$, then $P \Rightarrow P' \rightarrow_Y Q$ yields by induction $P \Rightarrow Q$. Then $M = YP \Rightarrow YQ = N'$.
- If Δ is the root redex YP' , then $N' = P'(YP')$, and we need only apply the Y -redex rule:

$$\frac{P \Rightarrow_p P' \quad YP \Rightarrow_p YP'}{YP \Rightarrow_p P'(YP')}$$

$YP \Rightarrow P'Q$: with $M = YP$ and $N = P'Q \xrightarrow{\Delta}_Y N'$. That is, the last derivation step looks as follows:

$$\frac{P \Rightarrow_p P' \quad YP \Rightarrow_p Q}{YP \Rightarrow_p P'Q}$$

Since P' cannot be Y itself, Δ must be in either P' or in Q . In the former case, we apply induction to $P \Rightarrow P' \rightarrow_Y P^*$ without changing the second hypothesis, so the conclusion of the rule becomes $YP \Rightarrow P^*Q$. In the latter case, we apply induction to $YP \Rightarrow Q \rightarrow_Y Q^*$ without changing the first hypothesis, so the conclusion becomes $YP \Rightarrow P'Q^*$, as desired. \square

Therefore, a Y -reduction sequence of arbitrary length can be turned into a single step of parallel reduction.

COROLLARY A.5

For all $M, N \in \Lambda_Y$, $M \rightarrow_Y N$ entails $M \Rightarrow N$.

Parallel reduction also satisfies the usual substitution property.

LEMMA A.6 (Substitution lemma for \Rightarrow).

For $M, M', N, N' \in \Lambda_Y$, we have

- $N \Rightarrow N'$ implies $M[N/x] \Rightarrow M[N'/x]$,
- $M \Rightarrow M'$ and $N \Rightarrow N'$ imply $M[N/x] \Rightarrow M'[N'/x]$.

PROOF OF LEMMA A.6.

- By the fact that \Rightarrow is a congruence.

(ii) By induction on $M \Rightarrow M'$. The only interesting case is the redex rule

$$\begin{array}{ll} M = (\lambda y.P)Q & P \Rightarrow P' \\ M' = P'[Q'/y] & Q \Rightarrow Q'. \end{array}$$

In this case

$$M[N/x] = (\lambda y.P[N/x])Q[N/x] \quad (\text{A.1})$$

$$M'[N'/x] = P'[Q'/y][N'/x] = P'[N'/x][Q'[N'/x]/y]. \quad (\text{A.2})$$

Note that the side condition $y \notin \text{FV}(N')$ needed for the application of the substitution lemma in (A.2) is inherited under $N \Rightarrow N'$ from the rules for capture-avoiding substitution in (A.1), where y is chosen implicitly to be such that $y \notin \text{FV}(N)$. By induction hypothesis, we have

$$\begin{array}{l} P[N/x] \Rightarrow P'[N'/x] \\ Q[N/x] \Rightarrow Q'[N'/x]. \end{array}$$

By applying the redex rule for \Rightarrow , we get

$$M[N/x] = (\lambda y.P[N/x])Q[N/x] \Rightarrow P'[N'/x][Q'[N'/x]/y] = M'[N'/x].$$

□

Acknowledgements

We are grateful to the anonymous reviewers whose suggestions helped improve and clarify this manuscript substantially.

References

- [1] R. M. Amadio and P.-L. Curien. *Domains and Lambda-Calculi*. Cambridge University Press, New York, NY, USA, 1998.
- [2] H. P. Barendregt. *The Lambda Calculus, Its Syntax and Semantics*, 2nd edn. Vol. 103 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, 1984.
- [3] I. Bethke. et al. *Lambda Calculus*, pp. 548–587. Vol. 55 of *Cambridge Tracts in Theoretical Computer Science*, 2003.
- [4] M. Bezem, J. W. Klop and R. de Vrijer. *Term Rewriting Systems—TeReSe*. Vol. 55 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2003.
- [5] A. Bucciarelli, A. Carraro, G. Favro and A. Salibra. Graph easy sets of mute lambda terms. *Theoretical Computer Science*, **629**, 51–63, 2016.
- [6] A. Carraro and A. Salibra. Ordered models of the lambda calculus. *Logical Methods in Computer Science*, **9**, 2013.
- [7] T. Coquand and H. Herbelin. A-translation and looping combinators in pure type systems. *Journal of Functional Programming*, **4**, 77–88, 1994.
- [8] J. Endrullis.. Private communication, 2011.
- [9] J. Endrullis, D. Hendriks and J. W. Klop. Modular construction of fixed point combinators and clocked Böhm trees. In *Proceedings of the 25th Annual IEEE Symposium on Logic in Computer Science, LICS 2010*, pp. 111–119. IEEE Computer Society, 2010.

- [10] J. Endrullis, D. Hendriks, J. W. Klop and A. Polonsky. Discriminating lambda-terms using clocked Böhm trees. *Logical Methods in Computer Science*, **10**, 2014.
- [11] J. Endrullis and R. C. de Vrijer. Reduction under substitution. In *Rewriting Techniques and Applications, 19th International Conference, RTA 2008, Hagenberg, Austria, July 15–17, 2008, Proceedings*, A. Voronkov ed., pp. 425–440. Vol. 5117 of *Lecture Notes in Computer Science*. Springer, 2008.
- [12] H. Geuvers and J. Verkoelen. *On fixed point and looping combinators in type theory*, 2009. <http://www.cs.ru.nl/herman/PUBS/TLCApaper.pdf>.
- [13] H. Geuvers and B. Werner. On the Church–Rosser property for expressive type systems and its consequences for their metatheoretic study. In *Proceedings Ninth Annual IEEE Symposium on Logic in Computer Science*, pp. 320–329. IEEE Computer Society, 1994.
- [14] J. M. E. Hyland. A syntactic characterization of the equality in some models for the lambda calculus. *Journal London Mathematical Society* (2), **12**, 361–370, 1975/76.
- [15] B. Intrigila. Non-existent Statman’s double fixedpoint combinator does not exist, indeed. *Information and Computation*, **137**, 35–40, 1997.
- [16] B. Intrigila. TLCA list, Problem #25: how many fixed points can a combinator have? 2000. <http://tlca.di.unito.it/optlca/>. (First raised in [17].)
- [17] B. Intrigila and E. Biasone. On the number of fixed points of a combinator in lambda calculus. *Mathematical Structures in Computer Science*, **10**, 595–615, 2000.
- [18] B. Intrigila and R. Statman. Lambda theories allowing terms with a finite number of fixed points. *Mathematical Structures in Computer Science*, 1–23, 2015.
- [19] J. W. Klop. New fixed point combinators from old. In *Reflections on Type Theory, lambda-Calculus, and the Mind*, pp. 197–210. Radboud University, 2007. (Essays dedicated to Henk Barendregt on the Occasion of his 60th Birthday.)
- [20] J. W. Klop. Personal communication, 2014.
- [21] A. Kurz, D. Petrisan, P. Severi and F. de Vries. Nominal coalgebraic data types with applications to lambda calculus. *Logical Methods in Computer Science*, **9**, 2013.
- [22] S. Lusin and A. Salibra. The lattice of lambda theories. *Journal of Logic and Computation*, **14**, 373–394, 2004.
- [23] G. Manzonetto and A. Salibra. From λ -calculus to universal algebra and back. In *Mathematical Foundations of Computer Science 2008: 33rd International Symposium, MFCS 2008, Torun, Poland, August 25–29, 2008, Proceedings*, E. Ochmanski and J. Tyszkiewicz, eds, pp. 479–490. Vol. 5162 of *Lecture Notes in Computer Science*. Springer, 2008.
- [24] L. Ong. Correspondence between operational and denotational semantics. In *Handbook of Logic in Computer Science*, S. Abramsky, D. Gabbay, and T. S. E. Maibaum, eds, vol. 4, pp. 269–365. Oxford University Press, 1995.
- [25] G. D. Plotkin. The lambda-calculus is ω -incomplete. *Journal of Symbolic Logic*, **39**, 313–317, 1974.
- [26] A. Polonsky. The range property fails for \mathcal{H} . *Journal of Symbolic Logic*, **77**, 1195–1210, 2012.
- [27] A. Salibra. Nonmodularity results for lambda calculus. *Fundamenta Informaticae*, **45**, 379–392, 2001.
- [28] P. Severi and F. de Vries. Order structures on Böhm-like models. In *Computer Science Logic: 19th International Workshop, CSL 2005*, C. L. Ong ed., pp. 103–118. Vol. 3634 of *Lecture Notes in Computer Science*. Springer, 2005.
- [29] R. Statman. Some examples of non-existent combinators. *Theoretical Computer Science*, **121**, 441–448, 1993.

- [30] R. Statman. On the lambda \mathcal{Y} calculus. *Annals of Pure and Applied Logic*, **130**, 325–337, 2004.
- [31] C. Wadsworth. The relation between computational and denotational properties for Scott's \mathcal{D}_∞ -models of the lambda-calculus. *SIAM Journal of Computing*, **5**, 488–521, 1976.

Received 00 Month 20xx